

INTELLIGIBILITY OF SPEECH PROCESSED THROUGH  
THE COCHLEA OF FETAL SHEEP *IN UTERO*

By

XINYAN HUANG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1999

Dedicated to my wife, Min Feng

## ACKNOWLEDGMENTS

First and foremost, I would like to express my sincerest appreciation and gratitude to my committee chairman and mentor, Dr. Kenneth Gerhardt, for his constant guidance, support, encouragement, and invaluable contribution to my professional and personal development. I would like to respectfully thank Dr. Robert Abrams for the constant support and encouragement he gave me regarding science, academics, and American culture.

Secondly, my thanks go to my committee members, Dr. Scott Griffiths, Dr. Francis Joseph Kemker, and Dr. Kyle Rarey, for their thoughtful suggestions and support. I would like to thank the faculty and staff in the Department of Communication Sciences and Disorders and in the Perinatology Research Laboratory for their valuable help during this study. I especially thank Mr. Rodney Housen for his computer programming assistance.

Finally, I wish to express my deepest appreciation and thanks to my wife. Without her love, patience, understanding, and continued support, this endeavor would have not been possible. My love and appreciation are imparted to my parents whose inspiration has kept me in constant pursuit of my dreams.

# TABLE OF CONTENTS

|   | page |
|---|------|
| ACKNOWLEDGMENTS .....   | iii  |
| LIST OF TABLES .....  | vi   |
| LIST OF FIGURES .....   | ix   |
| ABSTRACT .....  | xi   |
| CHAPTERS  |      |
| 1 INTRODUCTION .....  | 1    |
| 2 REVIEW OF LITERATURE .....                                      | 6    |
| Fetal Hearing .....   | 6    |
| Development of the Auditory System .....                          | 6    |
| Development of the Place Principle .....                          | 9    |
| Central Auditory System .....                                     | 13   |
| Fetal Behavioral Response to Sound .....                          | 14   |
| Fetal Sound Environment .....                                     | 16   |
| Intrauterine Background Noise .....                               | 16   |
| Sound Transmission into the Uterus .....                          | 19   |
| Fetal Sound Isolation .....                                       | 21   |
| Route of Sound Transmission into the Fetal Inner Ear .....        | 23   |
| Model of Fetal Hearing .....                                      | 25   |
| Intelligibility of Speech Sounds Recorded within the Uterus ..... | 27   |
| Fetal Auditory Experiences and Learning .....                     | 31   |
| Prenatal Effects of Sound Experience .....                        | 31   |
| Postnatal Effects of Prenatal Sound Experience .....              | 35   |
| Speech Perception .....   | 44   |
| Speech Perception in Infancy .....                                | 44   |
| Characteristic of Speech .....                                    | 45   |
| Intelligibility of Speech .....                                   | 48   |
| 3 MATERIALS AND METHODS .....                                     | 56   |
| Surgery .....   | 56   |

|   |     |
|---|-----|
| Recording Speech Stimuli .....                    | 58  |
| Perceptual Testing .....                          | 62  |
| Subjects .....                                    | 62  |
| Speech Stimuli .....                              | 62  |
| Procedures .....                                  | 64  |
| Data Analyses .....                               | 65  |
| Statistical Analyses .....                        | 65  |
| Information Analyses .....                        | 67  |
| Acoustic Analyses .....                           | 68  |
| 4 RESULTS AND DISCUSSION .....                    | 70  |
| Intelligibility .....                             | 70  |
| Consonant Feature Transmission .....              | 94  |
| Acoustic Analyses of Vowel Transmission .....     | 117 |
| 5 SUMMARY AND CONCLUSIONS .....                   | 153 |
| APPENDICES  |     |
| A SUBJECT RESPONSE SHEET .....                    | 158 |
| B RAW DATA FROM SUBJECT RESPONSE FORMS .....      | 161 |
| C RAW DATA FROM ACOUSTIC ANALYSES OF VOWELS ..... | 169 |
| REFERENCES .....                                  | 178 |
| BIOGRAPHICAL SKETCH .....                         | 190 |

## LIST OF TABLES

| <u>Table</u>  | <u>page</u> |
|---|-------------|
| 3-1 Perceptual tests .....  | 63          |
| 4-1 VCV stimulus intelligibility scores .....   | 76          |
| 4-2 CVC stimulus intelligibility scores .....   | 77          |
| 4-3 ANOVA summary table for VCV stimuli .....   | 78          |
| 4-4 <i>Post hoc</i> multiple comparisons (Newman-Keuls test) for VCV stimuli .....                    | 79          |
| 4-5 ANOVA summary table for CVC stimuli .....   | 80          |
| 4-6 <i>Post hoc</i> multiple comparisons (Newman-Keuls test) for CVC stimuli .....                    | 81          |
| 4-7 Consonant confusion matrix for male talker, recorded in air at 105 dB SPL .....                   | 95          |
| 4-8 Consonant confusion matrix for male talker, recorded in air at 95 dB SPL .....                    | 96          |
| 4-9 Consonant confusion matrix for male talker, recorded in the uterus at 105 dB SPL .....            | 97          |
| 4-10 Consonant confusion matrix for male talker, recorded in the uterus at 95 dB SPL .....            | 98          |
| 4-11 Consonant confusion matrix for male talker, recorded from CM-ex <i>utero</i> at 105 dB SPL ..... | 99          |
| 4-12 Consonant confusion matrix for male talker, recorded from CM-ex <i>utero</i> at 95 dB SPL .....  | 100         |
| 4-13 Consonant confusion matrix for male talker, recorded from CM-in <i>utero</i> at 105 dB SPL ..... | 101         |
| 4-14 Consonant confusion matrix for male talker, recorded from CM-in <i>utero</i> at 95 dB SPL .....  | 102         |

|      |  |     |
|------|--|-----|
| 4-15 | Consonant confusion matrix for female talker, recorded in air at 105 dB SPL ..   | 103 |
| 4-16 | Consonant confusion matrix for female talker, recorded in air at 95 dB SPL ....  | 104 |
| 4-17 | Consonant confusion matrix for female talker, recorded in the uterus at 105 dB SPL .....   | 105 |
| 4-18 | Consonant confusion matrix for female talker, recorded in the uterus at 95 dB SPL .....  | 106 |
| 4-19 | Consonant confusion matrix for female talker, recorded from CM-ex utero at 105 dB SPL .....  | 107 |
| 4-20 | Consonant confusion matrix for female talker, recorded from CM-ex utero at 95 dB SPL .....   | 108 |
| 4-21 | Consonant confusion matrix for female talker, recorded from CM-in utero at 105 dB SPL .....  | 109 |
| 4-22 | Consonant confusion matrix for female talker, recorded from CM-in utero at 95 dB SPL .....   | 110 |
| 4-23 | Conditional percentage of voicing, manner, and place information received (of bits sent) for each talker, recording location, and stimulus level condition for the nonsense syllables (VCV) .....  | 112 |
| 4-24 | Average fundamental frequencies ( $F_0$ ) and first three formant frequencies ( $F_1$ , $F_2$ , $F_3$ ) for five vowels produced by each talker and recorded in air .....  | 128 |
| 4-25 | Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ , $F_2$ , $F_3$ ) for vowel /i/ produced by each talker at different recording sites in the 105 dB condition ..... | 129 |
| 4-26 | Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ , $F_2$ , $F_3$ ) for vowel /I/ produced by each talker at different recording sites in the 105 dB condition ..... | 135 |
| 4-27 | Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ , $F_2$ , $F_3$ ) for vowel /e/ produced by each talker at different recording sites in the 105 dB condition ..... | 138 |

|      |  |     |
|------|--|-----|
| 4-28 | Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1, F_2, F_3$ ) for vowel /æ/ produced by each talker at different recording sites in the 105 dB condition ..... | 143 |
| 4-29 | Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1, F_2, F_3$ ) for vowel /ʌ/ produced by each talker at different recording sites in the 105 dB condition ..... | 146 |
| 4-30 | Summary of acoustic analyses of vowels .....   | 150 |



## LIST OF FIGURES

| <u>Figure</u>   | <u>page</u> |
|---|-------------|
| 3-1 Schematic drawing showing the animal and the setup of devices for stimulus generation, stimulus measurement, and recording in air, in the uterus, and from the fetal inner ear (cochlear microphonic) .....   | 59          |
| 3-2 Examples of CMs evoked by airborne pure tones at 0.5 and 2.0 kHz .....  | 61          |
| 3-3 The frequency responses of two types of earphones .....   | 66          |
| 4-1 Mean percent intelligibility of VCV nonsense stimuli spoken by a male and a female talker recorded in air, in the uterus, from the fetal CM <i>ex utero</i> , and from fetal CM <i>in utero</i> at two airborne stimulus levels .....                                   | 72          |
| 4-2 Mean percent intelligibility of CVC words spoken by a male and a female talker recorded in air, in the uterus, from the fetal CM <i>ex utero</i> , and from fetal CM <i>in utero</i> at two airborne stimulus levels .....  | 74          |
| 4-3 Mean percent intelligibility of VCV nonsense stimuli spoken by a male and a female talker recorded in air, in the uterus, from the fetal CM <i>ex utero</i> , and from fetal CM <i>in utero</i> when combining two airborne stimulus levels .....                       | 84          |
| 4-4 Mean percent intelligibility of CVC words spoken by a male and a female talker recorded in air, in the uterus, from the fetal CM <i>ex utero</i> , and from fetal CM <i>in utero</i> when combining two airborne stimulus levels .....                                  | 87          |
| 4-5 Conditional percentage of voicing, manner and place information received for a male (M) and a female (F) talker; in air (A), in the uterus (U), from the fetal CM <i>ex utero</i> (X), and from the fetal CM <i>in utero</i> (I); at 105 dB (H) and 95 dB (L) SPL ..... | 114         |
| 4-6 Spectrographic recordings of "Mark the word lash" at different recording conditions .....   | 119         |
| 4-7 Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ , $F_2$ , and $F_3$ ) for vowel /i/ produced by both talkers recorded at different locations at 105 dB SPL .....                                  | 131         |

|      |  |     |
|------|--|-----|
| 4-8  | Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ , $F_2$ , and $F_3$ ) for vowel /l/ produced by both talkers recorded at different locations at 105 dB SPL ..... | 137 |
| 4-9  | Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ , $F_2$ , and $F_3$ ) for vowel /e/ produced by both talkers recorded at different locations at 105 dB SPL ..... | 140 |
| 4-10 | Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ , $F_2$ , and $F_3$ ) for vowel /æ/ produced by both talkers recorded at different locations at 105 dB SPL ..... | 145 |
| 4-11 | Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ , $F_2$ , and $F_3$ ) for vowel /ʌ/ produced by both talkers recorded at different locations at 105 dB SPL ..... | 148 |

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

INTELLIGIBILITY OF SPEECH PROCESSED THROUGH  
THE COCHLEA OF FETAL SHEEP *IN UTERO*

By

Xinyan Huang

August 1999

Chairman: Kenneth J. Gerhardt  
Major Department: Communication Sciences and Disorders

The intelligibility of speech stimuli recorded from the fetal sheep inner ear (cochlear microphonic, CM) *in utero* was determined perceptually using a group of untrained judges. A fetus was prepared for acute recordings during a surgical procedure. Two separate lists, one of meaningful and one of nonmeaningful speech, were spoken by a male and a female talker, delivered through a loudspeaker to the side of a pregnant ewe, and recorded with an air microphone, a hydrophone placed inside the uterus, and an electrode secured to the round window of the fetus *in utero*. Perceptual test audio compact discs (CDs) generated from these recordings were played to 139 judges.

The intelligibility of the phonemes recorded in air was significantly greater than the intelligibility of these stimuli when recorded from within the uterus. The intelligibility of the phonemes recorded from CM *ex utero* was significantly greater than from CM *in utero*. Overall, male and female talker intelligibility scores recorded within the uterus averaged

91% and 85%, respectively. When recorded from the fetal CM *in utero*, intelligibility scores averaged 45% and 42% for the male and female talkers, respectively.

An analysis of the transmission of consonant feature information revealed that “voicing” is better transmitted into the uterus and into the fetal inner ear *in utero* than “manner” or “place.” Voicing information for the male, as well as manner and place information, was better preserved in the fetal inner ear *in utero* than for the female.

Spectral analyses of vowels showed that the fundamental frequency ( $F_0$ ) and the first three formants ( $F_1$ ,  $F_2$ , and  $F_3$ ) were well preserved in the uterus recordings for both talkers, but only  $F_0$ ,  $F_1$ , and  $F_2$  ( $< 2000$  Hz) were perceived in the fetal inner ear *in utero*. Only the lower frequency contents of vowels were present in fetal inner ear recordings.

This study demonstrated the presence of external speech signals in the fetal inner ear *in utero* and described the type of phonetic information that was detected at the fetal inner ear *in utero*.

## CHAPTER 1 INTRODUCTION

There is overwhelming evidence that the human fetus detects and responds to sound *in utero* (Querleu et al., 1989; Hepper, 1992; Lecanuet and Schaal, 1996). Studies in pregnant humans (Walker, Grimwade and Wood, 1971; Querleu et al., 1988a; Richards et al., 1992) and sheep (Armitage, Baldwin and Vince, 1980; Vince et al., 1982, 1985; Gerhardt, Abrams and Oliver, 1990) have shown the existence of a rich diversity of sound in the fetal environment, heavily dominated by the mother's voice and other internal noises and permeated by varied rhythmic and tonal sounds from the external environment. The human fetus has a well-developed hearing mechanism by the sixth month of gestation (Rubel, 1985a; Pujol and Uziel, 1988; Pujol, Lavigne-Rebillard and Uziel, 1990). During the last trimester, sound exposure may have a pronounced effect on fetal behavior and central nervous system maturation. Speech perception and voice recognition by the newborn may result directly from its prenatal experience (Fifer and Moon, 1988, 1995).

Linguistic theorists have proposed two alternative hypotheses regarding language development that infants upon birth are equipped with either a generalized auditory mechanism or a specialized speech-specific mechanism designed for perception of speech. Some theorists hold that human infants are born with a "speech module," a mechanism designed specifically for processing the complex and intricate acoustic

signals needed by humans to communicate with one another (Liberman, 1982; Fodor, 1983; Liberman and Mattingly, 1985; Wilkins and Wakefield, 1995; Fowler, 1996). An alternative theory of the neonate's initial state suggests that infants enter the world without specialized mechanisms dedicated to speech and language, but rather respond to speech using general sensory, motor, and cognitive abilities (Aslin, 1987; Kuhl, 1987, 1992; Jusczyk 1996; Ohala, 1996; Fitch, Miller and Tallal, 1997). Which theory, if either, applies to the human fetus is not known. What is known is that the fetus is beginning the dynamic process of acquiring the necessary skills for speech and language acquisition during prenatal life *in utero* (Querleu et al., 1989; Lecanuet, Granier-Deferre and Busnel, 1991; Lecanuet and Schaal, 1996).

The maternal voice is a naturally occurring and salient stimulus *in utero* that occurs during a crucial time period of fetal ontogeny (Querleu et al., 1988a; Benzaquen et al., 1990; Richards et al., 1992) in which several psychobiological systems, including the auditory system, are developing. The immediate effects of exposure to the mother's voice on the fetus may provide a way of tracking auditory system development, as well as measuring fetal ability to process sensory information (Fifer and Moon, 1988, 1994, 1995). Fetal auditory discrimination has also led to the hypothesis that prenatal experience with auditory stimulation is the precursor to postnatal linguistic development (Cooper and Aslin, 1989; Querleu et al., 1989; Ruben, 1992; Abrams, Gerhardt and Antonelli, 1998).

DeCasper and his colleagues (DeCasper and Fifer, 1980; DeCasper and Prescott, 1984) demonstrated that newborn infants preferred their mother's voice over that of other talkers. While this preference was assumed to be the product of *in utero* exposure to the

mother's voice and suggested that the fetus detected maternal vocalizations and retained memories of her speech patterns, it is not known what speech information actually reaches the fetal inner ear nor the extent to which the auditory system responds to externally generated speech. Querleu et al. (1988b) and more recently Griffiths et al. (1994) reported on the intelligibility of speech recorded with a hydrophone in the human (Querleu et al., 1988b) and sheep (Griffiths et al., 1994) uterus. In both studies, the recordings were played back to juries of normal listeners and speech intelligibility was calculated from their responses. The intelligibility of *in utero* recordings of speech was poorer than that of air recordings because the acoustic signature of human speech is modified by the abdominal wall, uterus, and amniotic fluids as it passes from air to the fetal head. The attenuation properties of the abdomen and uterus can be modeled as a low-pass filter with a high frequency cutoff at 250 Hz and a rejection rate of approximately 6 dB/octave (Gerhardt, Abrams and Oliver, 1990).

While the results of these studies reflect the perceptibility of the speech energies present in the amniotic fluid, they do not specify what speech energy might be present at the level of fetal inner ear. Measures of acoustic transmission to the fetal inner ear are quite limited at present (Gerhardt et al., 1992). Much work needs to be completed before conclusions can be drawn regarding what speech energies reach and are able to be perceived by the fetus.

The present experiment was designed to evaluate the intelligibility of speech produced through a loudspeaker and recorded with an electrode secured to the fetal sheep round window. The electrode recorded a bioelectric potential called the cochlear microphonic (CM). The CM is generated at the level of the hair cells and mimics the

input in amplitude and frequency (Gulick, Gescheider and Frisina, 1989). Recordings of the CM represent the time displacement patterns of the basilar membrane and reflect the initial response of the auditory periphery. The hypothesis is that speech is further degraded as it passes into the inner ear. Sheep were used in this study not only because sound attenuation characteristics of the abdominal contents of pregnant sheep are similar to those of pregnant women (Armitage, Baldwin and Vince, 1980; Querleu et al., 1988a; Gerhardt, Abrams and Oliver, 1990; Richards et al., 1992), but also because of the precocious hearing and the similarity of auditory sensitivity to humans. Sheep's hearing is only slightly poorer than that of humans for frequencies below about 8000 Hz (Wollack, 1963). The objective of this study was to determine what speech information was transmitted into the uterus and presented within the inner ear of the sheep fetus *in utero*.

The following hypotheses were tested:

1. The intelligibility of monosyllabic words and nonsense syllables will be reduced when recorded in the uterus compared to air.
2. The intelligibility of monosyllabic words and nonsense syllables will be reduced when recorded from the fetal inner ear *in utero* compared to uterus.
3. The intelligibility of a male talker will be greater than the intelligibility of a female talker when recorded in the uterus and from the fetal inner ear *in utero*.
4. Transmission into the uterus and fetal inner ear will be greater for voicing information than for manner and place information.



5. The transmission of voicing, manner, and place information will be better for males than for females when recorded in the uterus and from the inner ear of the fetus *in utero*.
6. Acoustic energy in the second and third formants of vowels measured in air for both male and female talkers will be reduced when recorded in the uterus, and will be reduced to the noise floor when recorded from the fetal inner ear *in utero*.

## CHAPTER 2 REVIEW OF LITERATURE

The human, unlike most mammalian species, is born with highly developed auditory sensitivity. By the 20th week of gestation, the structures of the peripheral auditory system, including the outer, middle, and inner ear, are anatomically like that of an adult, thus enabling the fetus to detect sounds during the last trimester of pregnancy (Rubel, 1985a; Pujol and Uziel, 1988; Pujol, Lavigne-Rebillard and Uziel, 1990). Responsiveness of the fetus to auditory stimuli begins during the 24th week of gestation (Birnholtz and Benacerraf, 1983; Shahidullah and Hepper 1993). Maturation of auditory processing capabilities takes place through prenatal and perinatal periods. An appreciation of the process of auditory development is important not only for an understanding of the normal auditory system, but also for an understanding of the impact of prenatal sound experience on the postnatal development, from structural, functional to behavioral development (Lecanuet and Schaal, 1996).

### Fetal Hearing

#### Development of the Auditory System

The earliest embryological signs of the human auditory apparatus are thickenings of the ectoderm on the sides of the head, bilaterally, called the auditory placodes. About

the 23rd day of gestational age (GA), each placode begins to invaginate to form the auditory pit, which then splits off from the overlying ectoderm to form an otocyst at the 30th day. At about 4 to 5 weeks, the otocyst divides into two parts, the vestibular portion and the cochlea. During the 8th through 11th week, the two and a half coils of the cochlea are attained. Complete maturation of sensory and supporting cells in the cochlea does not occur until the 20th week when the cochlea reaches adult size (Northern and Downs, 1991; Peck, 1994). Cytodifferentiation occurs during the 9th to 10th weeks within the cochlear duct, where there is a thickening of epithelium. From the 3rd to the 5th month, the thickening epithelium differentiates into the distinct receptor and supporting cells of the organ of Corti.

Comparing with that found in other mammals when the first responses to sound can be evoked, the human cochlea has achieved a functional stage by 20 weeks of gestation (Pujol and Uziel, 1988). At this time, the cochlea may have high thresholds and very poor discriminative properties. It is thus not possible to detect signs of cochlear activity using behavioral or electrophysiological methods, which explains why the first responses to acoustic stimulation can only be recorded a few weeks later (Starr et al. 1977; Birnholz and Benacerraf, 1983).

Rubel (1984) indicated that no single event triggers the onset of cochlear function. Many simultaneous and synchronous events contribute to the maturation of mechanical and neural properties. These events include thinning of the basilar membrane, formation of the inner spiral sulcus, maturation of the pillar cells, freeing of the inferior margin of the tectorial membrane, opening of the tunnel of Corti, formation of Nuel's spaces,

differentiation of the hair cells, establishment of mature cilia structure, and the maturation of synapses (Pujol and Hilding, 1973).

These final maturational events do not occur simultaneously throughout the length of the cochlea. There are two general developmental gradients in the differentiation and maturation of cochlea hair cells and their neural connections. The first is the classic basal to apical gradient, that at each maturation stage the mid-basal region develops first and spreads in both directions, with the apex maturing last. The second gradient is from inner hair cells (IHCs) to outer hair cells (OHCs); IHCs differentiate and develop first (Pujol and Uziel, 1988; Pujol, Lavigne-Rebillard and Lenoir, 1998). This does not necessarily imply that IHCs are the first to achieve all adult characteristics. For example, the completion of the ciliogenesis process occurs first at OHCs. Generally, synapse formation on IHCs occurs early and undergoes only minor modifications thereafter. The OHCs are initially surrounded by afferent terminals, which are gradually replaced by numerous efferents. Then the large calyciform efferent terminals form, typical of the mature cochlea.

Based on cat studies, the functional development of the auditory system is divided into three stages (Walsh and McGee, 1990). During the first stage, which is through the cats' first postnatal week and corresponds to the second trimester of human gestation, auditory responses can be elicited, but hearing thresholds are very high and well outside of the range of naturally occurring acoustic events. Response sensitivity does not significantly improve during this stage and the responsive frequency range is limited to low-frequency and mid-frequency sounds. During the second stage, in cats through the third postnatal week and in humans probably through the final trimester, rapid maturation

of auditory function takes place. Thresholds decrease substantially, the adult frequency response range is attained, and response duration is perceived. These changes are attributable in large part to cochlear maturation, and to a lesser extent to maturation of the central auditory system. During the final developmental stage, the remaining components within the auditory system mature slowly and myelination is complete. The adult characteristics for the cat are acquired during the second month after birth. However, further maturation of the human auditory system occurs after birth and continues for the next few years.

#### Development of the Place Principle

Young mammals do not respond initially to all of the frequencies to which they respond as adults. Generally, initial responses are elicited by low- or mid-frequency sounds. As development proceeds, responsiveness to both lower and higher frequencies increases. Responsiveness to the highest frequencies develops last (Rubel, 1978; Rubel, 1985a). However, cochlear differentiation occurs first in basal or mid-basal high-frequency regions, then spreads in both directions. The last part of the cochlea to undergo differentiation is the apical, low-frequency region (Rubel, 1978). A similar differentiation gradient also occurs in eighth-nerve ganglion cells and cochlear nuclei; regions receiving input from the basal, high-frequency region of the cochlea mature prior to the development of apical projection areas (Romand and Romand, 1982; Rubel, Smith and Miller, 1976; Schweitzer and Cant, 1984).

A paradox of cochlear development was pointed out by Rubel in 1978. During the early stages of hearing, the base or mid-basal region of the cochlea and the basal

representation areas of the central nervous system are the first to respond to sound. However, these areas are initially most sensitive to relatively low-frequency sound, even though this region of the cochlea has been tuned to being respond to high-frequency sound. With maturation of both mechanical and neural properties of the cochlea, the place code gradually shifts toward the apex until mature organization is achieved.

In an effort to understand more fully the mechanisms underlining this apparent paradox, Rubel and Ryals (1983) studied the position of hair cell damage produced by high-intensity pure tones of three different frequencies on three age groups of young chicks. The results showed that the position of maximum damage produced by each frequency shifted systematically toward the apex as a function of age. This experiment was carried out during the late stages of hearing development in the chick, corresponding to the perinatal or immediate postnatal periods in humans. On a related study, Lippe and Rubel (1983) evaluated the relationship between the location of neurons of the brainstem in chicks (nucleus magnocellularis and nucleus laminaris) and the frequency to which they were most sensitive. In both nuclei of the brainstem, embryonic neurons were most sensitive to tones 1-1.5 octaves below the frequencies that activate the same neurons one to two weeks after hatching. These two experiments provided support for the model of cochlear development offered by Rubel in 1978.

Later investigations, again in chicks, revealed some inconsistencies in the theory developed by Rubel (1978). The discrepancy between these studies may be attributed to developmental changes in middle-ear transfer function, the changes of the physical size of the basilar papilla, and temperature effects on frequency tuning (Rübsamen and Lippe, 1998). Currently, there are two alternative hypotheses for the development of the

cochlear frequency map in chicks. One theory suggests that frequency representation does not change developmentally. Another theory proposes that frequency representation shifts developmentally but that the shift is restricted to regions along the papilla that code mid- and high-frequency sounds, while low-frequency sounds are always represented at the apical location. Responses to mid-frequency sounds occur progressively more apically as the base becomes responsive to high-frequency sounds (Rübsamen and Lippe, 1998).

Dallos and his colleagues (Harris and Dallos, 1984; Yancey and Dallos, 1985; Arijmand, Harris and Dallos, 1988) studied the developmental change of the place code in gerbils. They reported that the cutoff frequency of the cochlear microphonic (CM) and the summing potential in the mid-basal turn (15 kHz location) increased about 1.5 to 2 octaves between the onset of sound evoked response on the 12th postnatal day when frequency representation becomes adultlike on the 21st postnatal days. But, the cutoff frequency of the CM at a second turn location (2.5 kHz) remains stable during development.

More direct evidence was provided by the finding that the characteristic frequencies of spiral ganglion neurons at a constant basal cochlear location increased up to 1.5 octaves between the second and third postnatal weeks (Echteler, Arjmand and Dallos, 1989). It has been uniformly reported that tonotopic organization in the mid- and high-frequency regions of the cochlea and central auditory nuclei changes during development. However, tonotopy in the cochlear apex and its central projection sites appeared to be developmentally stable (Rübsamen and Lippe, 1998). As a result of this new information, two updated explanations for the place code have been proposed. First,

the shifts in frequency code are attributed to maturational changes in the passive mechanical properties of the cochlea (Lippe and Rubel, 1985). Second, Romand (1987) proposed that the shifts in frequency organization should be attributed to maturational changes in cochlear active processes mediated by the outer hair cells. Both factors were examined by comparing tone-evoked distortion product otoacoustic emissions before and after an injection of furosemide in gerbils between 14 days old and adult (Mills, Norton and Rubel, 1994; Mills and Rubel, 1996). Results showed that increase in the passive base cutoff frequency rather than maturational changes in active processes accounts for the place code shift.

Currently, a revised model of the place code shift hypothesis for mammals, based on the evidence from developmental studies of central and peripheral frequency maps, is suggested. The entire length of the basilar membrane is capable of supporting a traveling wave at or very soon after the onset of hearing. Frequency representation in the cochlear apex is developmentally stable. From the very onset of hearing, the apex responds to its correct (adult) frequency, although the sensitivity and sharpness of tuning are reduced. In contrast, the more basal regions of the cochlea, mid- and high-frequency regions, undergo a shift in frequency organization such that each location becomes responsive to progressively higher frequencies in older animals. Shifts in the cochlear map result largely from maturational changes in the mechanical properties of the cochlear partition. The active mechanism also contributes to the shift in frequency organization (Rübsamen and Lippe, 1998).



### Central Auditory System

The development of the central auditory system and its relation to the maturation of the auditory periphery has been studied in animal models (Rubel, 1985a). Normal growth of central auditory neural elements requires an intact peripheral mechanism. However, initial stages of development of the auditory centers in the central nervous system are independent of peripheral regulation. The proliferation and migration of neurons in the central auditory system do not depend on the cochlea. The major pathways are established prior to or simultaneously with the development of peripheral function. Marty (1962) showed that in newborn kittens, the cortical evoked responses were elicited by electrical stimulation of the auditory nerve. The cochlea is immature at this time, and it is not possible to reliably evoke cortical responses to sound.

Following the establishment of functional connections between the periphery and the central nervous system, the continued maturation of neurons is highly dependent on the functional integrity of their afferents. Rubel and his colleagues (Rubel, Smith and Miller, 1976; Jackson, Hackett and Rubel, 1982) revealed that in chicks after the time when functional connections normally are established between the eighth nerve and the cochlear nucleus cells, the absence of peripheral innervation caused rapid and severe degeneration of the neurons. Abrams et al. (1987) demonstrated the impairment of glucose utilization in the auditory as well as nonauditory portions of the brain after cochlear ablation in fetal sheep.

### Fetal Behavioral Response to Sound

The human fetal auditory system is functional by the start of the third trimester (Birnholtz and Benacerraf, 1983). Although direct measurement of fetal hearing cannot be made by electrophysiological methods, indirect methods have been applied to measure fetal behavioral responses to sound stimuli. The most common approaches used to measure responsiveness to sound include the monitoring of fetal heart rate (Johansson, Wedenberg and Westen, 1964), fetal movement (Shahidullah and Hepper, 1994) and reflexive responses such as the auropalpebral reflex (Birnholtz and Benacerraf, 1983). Fetal movements in response to sound and to vibroacoustic stimulation or to both relate closely to the development of fetal audition (Gelman et. al, 1982; Hepper and Shahidullah, 1994a).

In 1983, Birnholtz and Benacerraf measured fetal responsiveness to an electronic artificial larynx (EAL) applied directly to the maternal abdomen. The auropalpebral reflex (blink-startle response) of the 236 fetuses tested from 16 to 32 weeks of gestation was monitored by ultrasonography. Reflexive eye movements were first elicited in some fetuses between 24 and 25 weeks of gestational age, and responses increased in frequency after 26 weeks. Consistent responses to EAL were observed after 28 weeks of pregnancy.

Shahidullah and Hepper (1993) examined the response of fetuses to a 110 dB SPL broadband air-borne stimulus (80-2000 Hz) at 15, 20 and 25 weeks of gestation. Using a response, which consists of a movement within 4.5 seconds of the onset of the stimulus, the investigators found that fetuses heard the noise at 25 weeks of gestation, but not earlier. However, when the stimulus was changed from a single pulse to a series of ten

pulses with two-second duration and ten-second inter-stimulus interval, a response was observed at 20 weeks of pregnancy. Thus, very early diffuse motor responses of slow latency were appeared as early as 20 weeks of gestation; by 25 weeks the response had become an immediate auditory startle response.

The auditory system of the fetus does not just begin to function uniformly across frequency. While the adult range of audibility is from 20 Hz to 20,000 Hz with greatest sensitivity in the 300 to 3000 Hz range, the fetus hears a much more limited range. Hepper and Shahidullah (1994b) examined the range of frequencies and intensity levels required to elicit human fetal movements as assessed with ultrasonography. Out of 450 fetuses involved in the study, only one demonstrated a response to a 500 Hz tone at 19 weeks gestational age. The range of frequencies to which the fetus responded expanded first to low frequencies, 100 Hz and 250 Hz, and then to high frequencies, 1000 Hz and 3000 Hz. By 27 weeks, 96% of the fetuses responded to tones at 100, 250 and 500 Hz, while none responded to frequencies at 1000 and 3000 Hz. It was not until weeks 29 (1000 Hz) and 31 (3000 Hz) that the fetuses responded to these tones. Between 33 and 35 weeks, the fetuses responded 100% of the time to presentations of 1000 and 3000 Hz. As gestation progressed from 19 to 37 weeks, the fetuses exhibited responsiveness to frequencies over a progressively wider frequency range. During this period, there was a significant decrease (20-30 dB) in the intensity level of stimulus required to elicit a response for all frequencies. This finding suggests that fetal hearing to pure tones becomes more sensitive as gestation proceeds.

The ability to discriminate frequency is fundamental for the interpretation of auditory information and for the development of speech perception and speech

production. Adults can detect changes of less than 2 Hz when the primary tone is between 100 Hz and 1000 Hz (Yost, 1994). The development of frequency discrimination in the human fetus was studied by Shahidullah and Hepper (1994) through the method habituation/dishabituation measurement. Ultrasound imaging was used to monitor fetal response to 250 and 500 Hz tones at 27 and 35 weeks gestation (N=48). They found that 35-week-old fetuses were capable of distinguishing between the two pure tones. However, fetuses at 27 weeks were not as likely to demonstrate this same discrimination.

Shahidullah and Hepper (1994b) also evaluated the abilities of 36 fetuses to differentiate between speech sounds. Fetuses at 27 and 35 weeks of age were exposed to a pair of pre-recorded syllables presented at 110 dB SPL through an earphone placed on the maternal abdomen. Half of the fetuses received /baba/ as their habituating stimuli and /bibi/ as their dishabituating stimulus and vice versa. Although all fetuses habituated, fewer stimuli were required for habituation for the 35-week-old fetuses than the 27-week-olds, and a greater number of the 35-week-old fetuses (17 of 18) demonstrated dishabituation compared to the younger ones (3 of 18). Thus, fetuses at thirty-five weeks possess the ability to discriminate among different phonemes.

### Fetal Sound Environment

#### Intrauterine Background Noise

The fetal sound environment is composed of a variety of internally generated noises, as well as many sounds originating from the environment of its mother. The once

held belief that the fetus develops in an environment devoid of external stimulation (Grimwarde, Walker and Wood, 1970) has been replaced by the fact that the fetus grows in the uterus filled with rich and diversified sounds originated inside and outside the mother (Gerhardt, 1989; Querleu et al., 1989).

The acoustic characteristics of internal noises and of external sounds that transmit into the uterus have been described in the human from various recording sites including inside the vagina (Bench, 1968), inside the cervix (Grimwarde, Walker and Wood, 1970), and inside the uterus following amniotomy (Querleu et al., 1988b; Benzaquen et al., 1990; Richards et al., 1992). These intrauterine sounds in humans were very similar to those recorded in pregnant sheep, via a chronically implanted hydrophone on the fetal head inside the uterus with an intact amniotic sac (Vince et al., 1982, 1985; Gerhardt, Abrams and Oliver, 1990).

Sounds generated inside the mother and present in the uterus are associated with maternal respiration (Vince et al., 1982; Gerhardt, Abrams and Oliver, 1990), maternal heartbeats (Walker, Grimwarde, and Wood, 1971; Querleu et al., 1988a), maternal intestinal activity (Vince et al., 1982; Gerhardt, Abrams and Oliver, 1990; Benzaquen et al., 1990), maternal physical movements (Vince et al., 1982; Gerhardt, Abrams and Oliver, 1990), and with placental and fetal circulation (Querleu et al., 1988a). These sounds provide a background or "noise floor" above which maternal vocalizations and externally generated sounds emerge (Vince et al., 1982, 1985; Querleu et al., 1988b; Gerhardt, Abrams and Oliver, 1990; Benzaquen et al., 1990; Richards et al., 1992).

In 1968, Bench measured the intrauterine noise floor at 72 dB SPL in a pregnant woman during labor. Three years later, Walker et al. (1971) reported an average intensity

of the background noise at 85 dB SPL (sound pressure level), with a peak at 95 dB SPL, which was associated with maternal heartbeats. However, the accuracy of these early studies was questioned by further studies using a hydrophone instead of a rubber-covered microphone previously used to measure the intrauterine sound level.

The use of a hydrophone represented an important technological improvement and provided more accurate data than was previously collected with air microphones. Studies in pregnant sheep (Vince et al., 1982; Gerhardt, Abrams and Oliver, 1990) and human (Querleu et al., 1988a; Benzaquen et al., 1990; Richards et al., 1992) showed that there is a quiet background with a muffled quality to sounds inside the uterus. Intrauterine sounds are predominately low frequency ( $< 100$  Hz) and reach 90 dB SPL (Querleu, Renard and Crépin, 1981; Vince et al., 1982; Gerhardt et al., 1990). Spectral levels decrease as frequency increases, and are as low as 40 dB for higher frequencies (Benzaquen et al., 1990; Gagnon, Benzaquen and Hunse, 1992). Gagnon et al. positioned a hydrophone in a pocket of fluid by the human fetal neck and measured sound pressure levels of 85 dB SPL at 12.5 Hz, decreasing to 60 dB for 100 Hz and less than 40 dB for 200 Hz and above. When measured in dBA, the human intrauterine sound level was only 28 dBA (Querleu et al., 1988a). Thus, for both humans and sheep, the noise floor tends to be dominated by low-frequency energy less than 100 Hz and can reach levels as high as 90 dB SPL.

Recently, Abrams et al. (1998) explored the origin of the intrauterine background noise in sheep under well-controlled laboratory conditions. The intrauterine noise level was measured before and after death of the ewe and fetus, and the average reduction in sound level postmortem approached 10-15 dB for frequencies below 100 Hz. The result

showed that sounds originating in the ewe and fetus contribute significantly to the low frequency (< 100 Hz) component of the background noise.

### Sound Transmission into the Uterus

Specifications of the amplitudes and frequency distributions of external sounds transmitted into the uterus have been well described in humans (Querleu et al., 1988a; Richards et al., 1992) and sheep (Armitage, Baldwin and Vince, 1980; Vince et al., 1982, 1985; Gerhardt, Abrams and Oliver, 1990). The attenuation of sound by the maternal abdominal wall, uterus and amniotic fluid is low in the low frequencies and increases in the high frequencies. In pregnant women, studied by Querleu et al. (1981), the attenuation is 2 dB at 250 Hz, 14 dB at 500 Hz, 20 dB at 1000 Hz and 26 dB at 2000 Hz. For high frequencies ranging from 3800 to above 18000 Hz, the attenuation is 20 to 40 dB (Querleu et al., 1988a). More recent results from Richards et al. (1992) showed that there was an average of 3.7 dB enhancement at 125 Hz, with progressively increasing attenuation up to 10.0 dB at 4000 Hz. Similar conclusions came from studies in sheep (Armitage, Baldwin and Vince, 1980; Vince et al., 1982, 1985; Gerhardt, Abrams and Oliver, 1990).

For frequencies below 250 Hz the reduction in sound pressure level through maternal tissue and fluids was less than 5 dB. Some enhancement of low-frequency sound pressures has been reported in both humans (Querleu et al., 1981; Richards et al., 1992) and sheep (Vince et al., 1982, 1985; Gerhardt, Abrams and Oliver, 1990). That is, the sound pressure in the amnion was greater than the sound pressure in air. Above 250 Hz, attenuation increased at a rate of about 6 dB per octave up to approximately 4000 Hz,

where the average attenuation was 20 to 25 dB. However, at 8000 Hz transmission loss was 15 dB (Gerhardt, Abrams and Oliver, 1990). These general findings have been refined and extended by Peters et al. (1993a, 1993b) who evaluated the transfer of airborne sounds across the abdominal wall of sheep as a function of frequency and intraabdominal location.

Peters et al. (1993a) studied the transmission of airborne sound into the abdomen of sheep over a wide frequency range (50-20,000 Hz). They found that mean attenuation varied from a high of 28 dB to a low of -3 dB. The greatest attenuation occurred for the frequencies between 5,000 and 12,500 Hz. Surprisingly, sound attenuation varied inversely as a function of stimulus level for low frequencies (50-125 Hz) and for high frequencies (7,000-20,000 Hz). At higher stimulus levels (110 dB SPL in air), attenuation was greater than the attenuation at lower stimulus levels (90 dB SPL). Thus, the 90 dB stimulus was more efficient than the 110 dB. In the middle frequency range (200-4,000 Hz), no effect of stimulus level was found.

In another study by Peters et al. (1993b), a hydrophone was positioned at each of 45 locations in a  $20 \times 20 \times 20$  array in the abdomen of five non-pregnant sheep post mortem. Isoattenuation contours within the abdomen were obtained. The sound pressure at different locations within the three-dimensional space of the sheep was highly variable. Low-frequency bands ( $< 250$  Hz) of noise revealed strong enhancement of sound pressure by up to 12 dB in the ventral part of the abdomen. For mid-frequencies (250-2000 Hz), attenuation reached as high as 20 dB. Attenuation for high frequencies ( $> 3150$  Hz) were somewhat less than for mid-frequencies and reached an upper limit of approximately 16 dB.



Over the frequency range from 250 to 4000 Hz, the abdomen can be characterized as a low-pass filter with high-frequency energy rejected at a rate of approximately 6 dB/octave (Gerhardt, Abrams and Oliver, 1990). Thus, external stimuli are shaped by the tissues and fluids of pregnancy before reaching the fetal head.

### Fetal Sound Isolation

It is known how much sound pressure is present at the fetal head. Now there is information about how much sound actually reaches the fetal inner ear (Gerhardt, et al. 1992). For the fetus *in utero*, external airborne sound energy must pass from the air medium to the fluid medium of the amnion before reaching the fetal inner ear. As sound energy changes medium, it is reduced because of the impedance difference at the air-tissue interface. The two quantities, pressure and particle velocity, are related and are dependent on the acoustic impedance of the medium. The acoustic impedance of water is much higher than that of air, for a given pressure disturbance, the particle velocity is much less by a factor of approximately 3600 ( $10 \log 3600 = 35.5$  dB) (Hawkins and Myrberg, 1983). Thus, equal pressure in air and fluid differ in sound energy by approximately 35 dB. One would assume that the sound pressure level required to produce a physiological response from the fetus would be approximately 35 dB greater than the sound pressure level in air necessary to produce the same response from the newborn (Gerhardt, 1990; Gerhardt, et al. 1992). Factors that determine how much *ex utero* sound reaches the inner ear of the fetus include the sound pressure attenuation through maternal tissue and fluid and the transformation of this pressure into basilar membrane displacement.

Gerhardt et al. (1992) studied the extent to which the fetal sheep *in utero* is isolated from sounds produced outside the mother. Inferences regarding sound transmission to the inner ear were made from cochlear microphonic (CM) input-output functions to stimuli with different frequency content. The CM, an alternating current generated by the hair cells of the inner ear, mimics the input signal in frequency and amplitude over a fairly wide range. As the signal amplitude increases, so does the amplitude of the CM. Cochlear microphonics recorded from the round window are sensitive indices of transmission characteristics of the middle ear. Thus, changes in the condition of the middle ear influence the amplitude of the CM. By comparing the sound pressure levels necessary to produce equal CM amplitude from the fetus *in utero*, and later, from the newborn lamb in the same sound field, estimates of fetal sound isolation can be made.

Cochlear microphonic input-output functions were recorded from *in utero* fetuses in response to one-third octave band noises from 125 to 2000 Hz and then again from the same animals after birth. The magnitude of fetal sound isolation was dependent upon stimulus frequency. For 125 Hz, sound isolation ranged from 6 to 17 dB, whereas for 2000 Hz fetal sound isolation ranged from 27 to 56 dB. The averages for each stimulus frequency were 11.1 dB for 125 Hz, 19.8 dB for 250 Hz, 35.3 dB for 500 Hz, 38.2 dB for 1000 Hz and 45.0 dB for 2000 Hz. Thus, for lower frequencies (< 500 Hz) the fetal auditory system appears to be sensitive to pressure variations produced by the stimulus originated from outside the mother.

### Route of Sound Transmission into the Fetal Inner Ear

Another factor that influences how airborne stimuli affects the fetus is related to the transmission of sound pressure from the fluid at the fetal head into the inner ear. Transmission is governed by the route that pressure variations take to reach the inner ear. The route of sound transmission postnatally is through the outer and middle ear system. Normal auditory function requires an air-filled middle ear cavity, an intact tympanic membrane, and functional hair cells and neural mechanism. In order to stimulate the hair cells of the inner ear, the movement of the stapes footplate in and out of the oval window creates hydraulic motion of the cochlear fluids, which causes basilar membrane displacement. However, in the fetus this route is likely to be rendered less efficient because the mechanical properties of the middle ear are highly dampened. The fetal middle ear and external ear canal are filled with amniotic fluid, which decreases the mechanical advantage of the middle ear. In addition, sound pressure may be present with the same phase at the oval window and round window. The lack of a phase difference, as well as the lack of a middle ear amplifier, may substantially decrease basilar membrane displacement and therefore cause a decrease in hearing sensitivity.

Two hypotheses have been proposed that describe the route that exogenous sounds take to reach the fetal cochlea. It has been suggested that acoustic stimuli in the fetal environment pass easily through the fluid-filled external auditory canal and middle ear system to the inner ear (Rubel, 1985b; Querleu et al., 1989). The impedance of inner ear fluids is similar to that of amniotic fluid, thus, little acoustic energy is lost due to an impedance mismatch (Querleu et al., 1989).

Hearing via bone conduction is a second alternative. Researchers have shown that the contribution of the external auditory meatus to auditory sensitivity in underwater divers is negligible (Hollien and Feinstein, 1975). By comparing the ability of a diver to hear under different conditions while in water, bone conduction has been shown to be much more effective in transmitting underwater sound energy. Similarly, fetal hearing occurs in a fluid environment and sound transmission may be through bone conduction as well.

Gerhardt, et al. (1996) compared the effectiveness of the two routes of sound transmission (outer and middle ear vs. bone conduction) by recording CM amplitudes from fetus sheep *in utero* in response to airborne sounds. CM input-output functions were obtained from the fetus *in utero* during three different conditions: uncovered fetal head, covered entire fetal head, and covered fetal head with exposed pinna and ear canal.

Results showed that when the fetal head was covered with sound attenuating material, even though the pinna and ear canal remain uncovered, sound levels necessary to evoke a response were greater than those necessary to evoke the same response from the fetus with its head uncovered. This fact revealed that acoustic energy in amniotic fluid reaches the fetal inner ear through a bone conduction route. External sounds transmitted into uterus stimulate the inner ear by vibrating fetal skull directly, which in turn results in the basilar membrane displacement. Thus, more sound energy is necessary to vibrate the skull to stimulate hair cell by bone conduction than by air conduction.

### Model of Fetal Hearing

Gerhardt and Abrams (1996) proposed a model of fetal hearing that considers what sounds are present in the environment of the fetus and to what extent these sounds can be detected. The model includes information regarding intrauterine background noise, sound transmission through the tissues and fluids associated with pregnancy and sound transmission through the fetal skull into the inner ear.

For the fetus to detect a signal from outside the mother, extrinsic sounds have to exceed the ambient sound level *in utero*. The internal noise floor of the mother is dominated by low-frequency energy produced by respiration, intestinal function, cardiovascular system, and maternal movements. Spectral levels decrease as frequency increases, and are 60 dB for 100 Hz and lower than 40 dB for 200 Hz and above. Presumably, the ability of the fetus to detect exogenous sounds will be dependent in part on the spectrum level of the noise floor because of masking effects. As expected, high-frequency sound pressures would be reduced by about 20 dB. The attenuation of low-frequency sounds by the abdominal wall, uterus and fluids surrounding the fetal head is quite small and in some cases enhancement of sound pressure of about 5 dB has been noted. Between 250 and 4000 Hz, sound pressure levels drop at a rate of 6 dB/octave. At 4000 Hz, maximum attenuation is approximately 20 dB. At frequencies higher than 4000 Hz, the attenuation is reduced to less than 20 dB.

Sound pressures at the fetal head create compressive forces through bone conduction that result in displacements of the basilar membrane thereby producing a CM. For 125 and 250 Hz, an airborne signal would be reduced by 10-20 dB in its passage to the fetal inner ear over what would be expected to reach the inner ear of the organism in

air. For 500 through 2000 Hz, the signal would be reduced by 40-45 dB. For frequencies in this range, the fetus is indeed buffered from sounds in the environment surrounding its mother probably because of limited function of the ossicular chain. However, for low-frequency sounds, the fetus is not well isolated. Low-frequency stimuli reach the inner ear of the fetus with far greater amplitudes than high-frequency stimuli. Interestingly, the development of the inner ear is such that low-frequency stimuli are detected before high-frequency stimuli. If the development of normal function is dependent on external stimulation, then the developmental pattern of the auditory system provides a mechanism to ensure each neuronal regions receive adequate stimulation from the environment (Rubel, 1984).

The fetus *in utero* will detect speech, but probably only the low-frequency components less than 500 Hz, and only when the airborne signal exceeds about 60 dB SPL. If it is less than that, the signal could be masked by internal noises. It is predicted that the human fetus could detect speech at conversational levels (65-75 dB SPL), but would not be able to discriminate many of the speech sounds with high-frequency components. Likewise, if music was played to the mother at comfortable listening levels, the temporal characteristics of music, rhythms, could be sensed by the fetus, but the high-frequency overtones would not be of sufficient amplitude to be detected (Abrams et al., 1998). Simply put, the fetus would be stimulated by music with the "bass" register turned up and the "treble" register turned down. This information may relate to *in utero* development of speech and language, to musical preferences and to subsequent cognitive development.

### Intelligibility of Speech Sounds Recorded within the Uterus

Speech produced during normal conversation is approximately 70 dB SPL and is comprised of acoustic energy primarily between 200 and 3000 Hz. The average fundamental frequency of an adult is 125 Hz for male's voice, and is 220 Hz for female's voice. Speech becomes unintelligible when the background noise in the speech-frequency range exceeds the level of the message by approximately 10 dB.

There are many factors that determine how well a fetus will hear sounds from outside its mother. These factors include: the frequency content and level of the internal noise floor; the attenuation of external signals provided by the tissues and fluids surrounding the fetal head; sound transmission into the fetal inner ear; and the sensitivity of the auditory system at the time of sound stimulation.

As a result of experimental work, the characteristics of the intrauterine sound environment are now fairly well understood. Studies in sheep (Vince et al., 1982, 1985; Gerhardt, Abrams and Oliver, 1990) and in humans (Querleu et al., 1988a; Benzaquen et al., 1990; Richards et al., 1992) have shown that the mother's voice and speech sounds from outside the mother transmit easily into the uterus with little attenuation, and form part of the intrauterine sound environment. Vince et al. (1982, 1985) implanted a hydrophone inside the amniotic sac of pregnant ewes, and obtained long-term recordings. They showed that the sound of maternal vocalizations forms a prominent part of the intrauterine sound environment, and is louder inside the uterus than outside. Gerhardt et al. (1990) also noted that when listening to the internal recordings from sheep, conversations were recognized between experimenters with normal vocal effort 3 feet from the ewe. Speech was muffled and intelligibility was poor, however, pitch,

intonation, and rhythm were quite clear. These findings are in accordance with data provided by human studies. Querleu et al. (1988b) presented various human voices through a loudspeaker to pregnant women and recorded the speech with a hydrophone in the uterus. The voice included mother talking directly, the mother's voice recorded on tape and playback, and the recorded voices of other women and men. All types of recorded voices (presented at 60 dBA) emerged above the basal noise floor (28 dBA) by +8 to +12 dB. The mother's voice recorded directly was 24 dB greater than the noise floor. The intensity of the maternal voice transmitted to the uterine cavity was greater than that of outside voices. Moreover, it was also transmitted to fetus more often than any other voices. In 1990, Benzaquen et al. reported that maternal vocalization was easily recorded *in utero* in ten pregnant women tested in the study. The sound spectrum produced by pronouncing the words of "99" was characterized by peak intensity of 70 to 75 dB SPL at 200 to 250 Hz and was approximately 20 dB above the intrauterine background noise at those frequencies.

Richards et al. (1992) studied the transmission of speech into the uterus. Intrauterine sound pressure levels of the mother's voice were enhanced by an average of 5.2 dB in the low-frequency range, whereas external male and female voices were attenuated by 2.1 and 3.2 dB, respectively. However, these studies only provided the information about the existence of speech sound in the intrauterine sound environment. The understandability of speech recorded from within the uterus is another critical issue for our understanding of early speech and language development. Fetal identification of its mother's voice and its ability to form memories of early exposure to speech are in part dependent on the intelligibility of the speech message.



Currently, two published studies address the perceptibility of speech recorded from inside the uterus. Querleu et al. (1988b) recorded the voices of five pregnant women and voices of other male and female talkers with a modified microphone positioned by the head of the fetus. Six listeners were able to recognize about 30% of the 3120 French phonemes. No significant difference was noted between the male and female voice, and the mother's voice was not better perceived although more intense. The recognition of vowels was correlated with their second formant. The intonation patterns, which frequencies were ranging from 100 to 1000 Hz, were perfectly well discriminated compared to linguistic meaning.

In a more recent study conducted by Griffiths et al. (1994), a panel of over 100 untrained individuals judged the intelligibility of speech recorded *in utero* from a pregnant ewe. Two separate word lists, one of meaningful and one of non-meaningful speech stimuli were delivered to the side of the ewe through a loudspeaker and were simultaneously recorded with an air microphone located 15 cm from the flank and with a hydrophone previously sutured to the neck of the fetus. Perceptual test tapes generated from these recordings were played to 102 judges. Intelligibility was influenced by three factors: transducer site (maternal flank or *in utero*); gender of the talker (male or female); and intensity level (65, 75 or 85 dB). For recordings made at the maternal flank, there was no significant difference between male and female talkers. Intelligibility scores increased with increased stimulus level for talkers and at both recording sites. However, intelligibility scores were significantly lower for females than for males when the recordings were made *in utero*.

An analysis of the feature information from recordings inside and outside the uterus showed that voicing information is better transmitted *in utero* than place or manner information. "Voicing" refers to the presence or absence of vocal fold vibrations (e.g., /s/ vs. /z/), "place" of articulation refers to the location of the major air-flow constriction during production (e.g., bilabial vs. alveolar), and "manner" refers to the way the speech sound is produced (e.g., plosive vs. glide).

Miller and Nicely (1955) reported that low-pass filtering of speech signals resulted in a greater loss of manner and place information than of voicing information. They concluded that the higher frequency information in the speech signal is critical for accurate identification of manner and place of articulation. The findings of Griffiths et al. (1994) are consistent with those of Miller and Nicely (1955) in that transmission into the uterus can be modeled as a low-pass filter. The poorer *in utero* reception of place and manner information is associated with the greater high frequency attenuation.

Voicing information from the male talker, which is carried by low-frequency energy, was largely preserved *in utero*. The judges evaluated the male talker's voice equally well regardless of transducer site. Speech of the female talker carried less well into the uterus. The fundamental frequency of the female talker was higher than that of the male talker. Thus, it is understandable that voicing information from the male would carry better into the uterus than that from the female.

Male and female talker intelligibility scores averaged approximately 55% and 34%, respectively, when recorded from within the uterus. Although these results reflect the perceptibility of the speech energies present in the amniotic fluid, they do not specify what speech energy might be present at the fetal inner ear. Measures of acoustic

transmission to the fetal inner ear are quite limited at present. Much work needs to be completed before conclusions can be drawn regarding what speech energies reach and are able to be perceived by the fetus.

### Fetal Auditory Experiences and Learning

During the last trimester, the human fetus, with a well-developed hearing mechanism, is exposed to a large variety of simple and complex sounds. Prolonged exposure, for several weeks or even months, to external and maternal sounds may have several consequences to the fetus at structural, functional, and behavioral levels. Prenatal activation of the auditory system may contribute to normal development of peripheral structures and central connections, as well as maintenance of anatomic and functional integrity during prenatal maturation. On a more general level, fetal auditory stimulation may contribute to the formation of auditory perceptual abilities, and to the organization of the newborn's preferences for a particular acoustical signal (Lecanuet and Schaal, 1996).

### Prenatal Effects of Sound Experience

Human fetal responsiveness to intense acoustical stimulation has been studied only in the past two decades. Fetuses are not only responsive to intense stimulation, they also display differential auditory responses as a function of the characteristics of the stimulus. When acoustic or vibroacoustic stimuli are above 110 dB SPL, fetuses display heart rate accelerations and motor-startle movement responses. Below 100 dB SPL, no reliable movement responses can be recorded, but fetuses display small, transient heart-rate decelerations rather than heart-rate accelerations (Lecanuet, Granier-Deferre and

Busnel, 1989, 1995). The heart-rate acceleration changes to auditory stimulation are typically associated with so-called "startling" or defensive response, while deceleration changes are "orienting" or attentive response (Berg and Berg, 1987).

Experiments have shown that repetition at a short interval (every 3-4 seconds) of a 92 to 95 dB SPL acoustic stimulus led to the disappearance of a cardiac deceleration response that had been induced by the first presentation of the stimulus, indicating an habituation (Lecanuet et al., 1992). Habituation is defined as the decrement in response after repeated presentation of a stimulus. Habituation is essential for the efficient functioning and survival of the organism, enabling it to ignore familiar stimuli and attend to new stimuli. Habituation represents one of the simplest yet most essential learning processes the individual possesses, and underlies much of our functioning and development (Hepper, 1992). Using a classical habituation / dishabituation procedure, Kisilevsky and Muir (1991) obtained a significant decrement of both fetal cardiac acceleration and movement responses to a complex noise (at 110 dB SPL), followed by a recovery of these responses when triggered by a novel vibroacoustic stimulus. The fetuses were between 37 and 42 weeks gestation during the experiment. Habituation *in utero* relates not only to the reception of the sensory message, but also its integration at lower levels of the central nervous system. Therefore, the fetus *in utero* is capable of learning (Querleu et al., 1989).

Lecanuet et al. (1989, 1993) studied the auditory discriminative capacities of the near-term fetus by using habituation/dishabituation of heart-rate deceleration responses. In one study (Lecanuet, Granier-Deferre and Busnel, 1989), fetuses at 35 to 38 weeks gestation displayed a transit heart-rate deceleration response when they were exposed to

the repeated presentation (every 3.5 second) of a pair of French syllables: /ba/ and /bi/ or /bi/ and /ba/, spoken by a female talker at 95 dB SPL. Reversing the order of the paired syllables after 16 presentations also reliably induced the same type of response. This was observed in 15/19 fetuses in the BABI/BIBA condition and in 10/14 fetuses in the BIBA/BABI condition. Response recovery suggested that fetuses discriminated between the two stimuli. The discrimination that occurred may have been performed on the basis of a perceptual difference in loudness (intensity) between the /ba/ and /bi/, since the equalization of these syllables was presented with SPL, not hearing level. This intensity adjustment makes /bi/ louder than /ba/ for adult listeners. Similarly, Shahidullah and Hepper (1994) found that fetuses at 35 weeks gestation had the ability to discriminate between /baba/ and /bibi/.

In another experiment (Lecanuet et al., 1993), the ability of near-term fetuses to discriminate different speakers producing the same sentence was studied. The heart-rate responses of fetuses between 36 to 39 weeks gestation were recorded before, during and after stimulation to the sentence 'Dick a du bon thé' (Dick has some good tea). The sentence was spoken by either a male talker (minimum fundamental frequency  $F_0 = 83$  Hz) or a female talker (minimum  $F_0 = 165$  Hz) and delivered through a loudspeaker 20 cm above the mother's abdomen at the same level (90-95 dB SPL). The fetuses were exposed to the first voice presentation (male or female) and followed by the other voice or the same voice (control condition) after fetal heart-rate response returned to baseline. The results demonstrated that in the first 10 s after presentation of the initial voice, the voice (male or female) induced a high and similar proportion of heart rate deceleration changes (77% to the male voice, 66% to the female voice) compared to a group of non-

stimulated subjects (9% of deceleration and 46% of acceleration). Within the first 10 s following the voice change, 69% of the fetuses exposed to the other voice displayed a heart-rate deceleration response, whereas 43% of the fetuses in the control condition displayed heart-rate acceleration change. The authors pointed out that near-term fetuses might perceive a difference between the voice characteristics of two speakers, at least when they are highly contrasted for  $F_0$  and timbre. The results cannot be generalized for all male and female voices or for all speakers since voices with extremely low  $F_0$  were used in the study (Lecanuet, Granier-Deferre and Busnel, 1995; Lecanuet, 1996).

Hepper et al. (1993) studied the ability of fetuses to discriminate between a strange female's voice and the mother's voice by measurement of the number of fetal movements during a 2-minute speech presentation. The results showed that fetuses at 36 weeks gestation did not discriminate between their mother's voice and that of a stranger, when tape recordings were played to them via an air-coupled loudspeaker placed on the abdomen. However, the fetuses were able to discriminate between their mother's voice recorded on tape and played to them over the loudspeaker and the mother's voice produced naturally; less movements were noted in response to the mother's direct speaking voice when compared to a tape recording of her voice. According to the authors, discrimination may be due to the presence of internally transmitted components of speech which the fetus perceives when the mother is speaking, but that are not present when the tape recording of the mother's voice is played.

The possibility of prenatal recognition of a familiar child's rhyme was studied by DeCasper et al. (1994). Seventeen pregnant women recited a child's rhyme aloud three times a day from their 33rd to 37th week of pregnancy. Fetal heart-rate response was

used to assess differential fetal responsiveness to the target rhyme versus a novel rhyme. During the 37th week of gestation, each fetus was stimulated to one rhyme for 30 seconds through a loudspeaker placed over the mother's abdomen. The first rhyme was followed by 75 s of silence and then the other rhyme was presented for 30 s. Stimulus level for both rhymes was set at 80-82 dB SPL. Care was taken during fetal testing to keep the mother unaware of which rhyme was being presented so that she could not inadvertently cue her fetus. The results showed that fetal heart rates significantly decreased from prestimulus levels when the target rhyme was presented and significantly increased over prestimulus levels when the novel rhyme was presented, regardless of presentation order. This differential heart-rate change implied that the fetus discriminated the two rhymes. Moreover, since these rhymes were counterbalanced across fetuses, the different patterns of heart-rate responds could not be attributed to any unique acoustic attributes of one rhyme.

There is now a growing body of data showing that fetuses perceive acoustical stimuli. Near-term fetuses can discriminate between two complex stimuli (such as syllables), between two speech passages, and they are able to learn. Such a competence may be partly a consequence of fetal familiarization to speech sounds.

#### Postnatal Effects of Prenatal Sound Experience

Prenatal auditory experience may result in general and / or specific learning effects that are evidenced in postnatal life. Stimuli familiar to the fetus may selectively soothe the baby after birth or may elicit orienting responses during quiet states. Familiar stimuli are more alerting than unfamiliar ones. It is well documented that prenatal

auditory experience plays a major role in the development of human newborn auditory preferences and capabilities (Fifer, 1987; Leanuet, 1996).

It has been shown that maternal heartbeat (Salk, 1962) and recordings of intrauterine noises (Rosner and Doherty, 1979) can calm a restless baby and serves as a potent reinforcer during operant conditioning nonnutritive sucking procedures (DeCasper and Sigafos, 1983). Indeed, intrauterine cardiac rhythms are potent reinforcers for 2- to 3-day-old newborns, a finding that suggests that prenatal auditory experience affects postnatal behavior.

Nonnutritive sucking procedures made it possible to objectify newborn's discriminative abilities and to test the newborn's preference for a given stimulus. The human voice, especially that of its mother, is likely to have increased salience for the fetus relative to other auditory stimuli. Mother's voice in the fetal sound environment differs from other sounds in its intensity, variability, and other multimodal characteristics. Mother's voice has been reported to be the most intense acoustic signal measured in the amniotic environment (Querleu et al., 1988a; Benzaquen et al., 1990; Richards et al., 1992). The nature of the maternal voice may promote greater fetal responsiveness to mother's voice than any other prenatal sound. The earliest evidence for differential responsiveness to maternal voice came from work with older infants (Mills and Meluish, 1974). The experiments demonstrated a differential sensitivity to the maternal voice in 20- to 30-day-old infants. The amount of time spent sucking and number of sucks per minute were increased after a brief presentation of his/her mother's voice. In a later study using 1-month-old infants (Mehler et al., 1978), sucks were reinforced with either a mother's or a stranger's voice, intonated or monotone. A significant increase in sucking



was only observed when mother's voice was normally intonated. The role of intonation in recognition of the mother's voice was suggested. Although these procedures clearly demonstrate that infants respond differentially to their mother's normal voice, the differences in responding do not necessarily indicate a preference for her voice (Fifer, 1987).

The study by DeCasper and Fifer (1980), using two different nonnutritive sucking procedures, was the first to provide direct experimental evidence that neonates prefer their mother's voice. Using a temporal discrimination procedure, 2- to 3-day-old infants were observed for a 5-minute baseline period in which nonrewarded sucks on a nonnutritive nipple were recorded. The median time of the interburst intervals (IBIs) was calculated and used to set the contingency for the testing. For 5 of the 10 infants tested, sucking bursts that ended IBIs shorter than the baseline median IBI (mIBI) turned on a tape recording of the infant's mother reading a children's story. Whereas sucking bursts that ended IBIs equal to or longer than the mIBI turned on a tape recording of another infant's mother reading the same story. For the other five infants, the IBI/story contingency was reversed. The results showed that 8 of the 10 infants shifted their overall medians significantly in the direction necessary to turn on the recording of its mother's voice. Also, the infants turned on the recording of their mother's voice more often and for a longer total period of time than the unfamiliar female voice.

In the second procedure, which involved a signal discrimination paradigm, the presence or absence of a 4-s 400 Hz tone signaled the availability of the different voices, and the voices remained on for the duration of the sucking burst. For 8 of the 16 infants tested, sucking on the nipple during the tone resulted in the cessation of the tone and

turned on a recording of their own mother's voice reading a children's story, whereas sucking during silence turned on a recording of another woman reading the same story. For the other eight infants, the signal/story contingency was reversed. Again, evidence of newborns' preference for their own mother's voice was obtained. Infants showed a significantly greater probability of sucking during the signal (tone or silence) that led to the presentation of the maternal voice recording.

Since it is possible that preference for the mother's voice could be generated very fast by the newborn's initial postnatal contact with the mother, several subsequent studies have attempted to rule out the effect of postnatal auditory experience. Fifer (1987) failed to find any evidence that preference in newborns for maternal voice was related to either postnatal age (1- vs. 3-day-olds) or method of feeding (bottle-fed vs. breast-fed). Another study showed that 2-day-old newborns did not prefer its father's voice to that of another male's voice, even though these newborns had 4 to 10 hours of postnatal contact with their fathers (DeCasper and Prescott, 1984). This study also determined that the absence of a preference for the paternal voice was not due to the inability of newborns to discriminate between pairs of male voices. Furthermore, the authors compared the preference between an airborne version of those mother's voice and their "intrauterine", low-pass filtered version. Using tone/silence discriminative responding procedures, 2- to 3-day-old infants were given a choice of hearing their mother's voice (or other female's voice) either unfiltered or low-pass filtered at 1000 Hz (Spence and DeCasper, 1987). Infants showed no preference for either the unfiltered or low-pass filtered version of their mother's voice, whereas infants preferred the unfiltered version of the nonmaternal voice to the filtered nonmaternal voice. According to the authors, since there is apparently little

prenatal experience with the low-frequency features of other female voices, but considerable postnatal experience with their full spectral characteristics, the newborns preferred the more familiar version of the female stranger's voice. In contrast, both the filtered and unfiltered versions of maternal voice contained the necessary low-frequency features for maternal voice recognition, so the infants showed no preference.

Finally, Fifer and Moon (1989), using a modified version of the "intrauterine" mother's voice mixed or not mixed with maternal cardiovascular sounds, found that 2-day-old newborns preferred a low-pass filtered version of the maternal voice to an unfiltered version when 500 Hz was the cutoff frequency. Therefore, it is possible that the infants in the previous study (Spence and DeCasper, 1987) did not show a preference for the filtered maternal voice because it was more similar to their postnatal rather than their prenatal experience with the maternal voice. Newborns' prenatal familiarity with maternal voice may explain the findings by Hepper et al. (1993). Using an analysis of fetal movements, Hepper et al. demonstrated that 2- to 4-day-old newborns discriminated normal speech from "motherese" speech of their mothers' voice, but not between normal intonated and one of "motherese" of a strange female's voice. Newborns, however, discriminated the maternal voice from a strange female voice.

Taken together, these results suggest that prenatal auditory experience determines at least some of the infant's early auditory preferences. This prenatal effect was demonstrated more directly by the study conducted by DeCasper and Spence (1986). Sixteen pregnant women recited one of the three children's stories aloud twice each day during the final 6 weeks of their pregnancies. After birth, the newborns (average age of 55.8 hours) were tested using the nonnutritive IBI contingent sucking procedure. For

eight of the infants in the prenatal group, sucking bursts following IBIs < mIBI turned on a recording of a woman (either the infant's own mother or the mother of another infant) reading the story that the infant's mother had read while pregnant. Sucking bursts which followed IBIs  $\geq$  mIBI turned on a recording of that same woman reading a novel story. For the other eight infants in the prenatal group, the IBI/story contingency was reversed. Additionally, a control group (12 infants) was tested under the same conditions except that these infants had no experience with any of three stories. The results showed that regardless of which story the mothers had recited while pregnant and regardless of the IBI/story contingency, the newborns in the prenatal group were more likely to suck after IBIs required to turn on the familiar story, the one they had heard prenatally, whereas infants in the control group showed no systematic change in their sucking pattern from baseline. Moreover, these preferences for one of three stories were not dependent on the specific voice of the storyteller. This result showed that the induction of a preference for a story (speech passage) generalized from maternal to nonmaternal voice. It implies that the newborn retains two different kinds of acoustic information from prenatal experience: information about specific characteristics of the mother's voice (perhaps fundamental frequency) and more general characteristics that are not necessarily mother-specific, such as intonation contours and / or temporal characteristics.

These studies provide strong evidence that the late-term human fetus is able to process some aspects of vocal stimulation presented by the mother and retain some of that information for at least several days after birth. It remains unclear, however, which specific aspects of prenatal auditory stimulation were responsible for postnatal auditory preferences.

Because external low-frequency sound is transmitted into the uterus with little attenuation and because high-frequency sound is attenuated, the fetus can only detect the low-frequency components of passage presented by the mother. It appears that these newborns could not merely depend on segmental information (phonetic components of speech, i.e., the specific consonants and vowels making up the words), which they experienced prenatally, as the basis for their postnatal recognition, since segmental information is carried by those frequencies that appear to be most attenuated *in utero* (frequencies above 1000 Hz). In contrast, the suprasegmental information (intonation, frequency variation, stress, and rhythm) contained in the maternal voice and in the stories recited by the mother is available to the fetus with very little attenuation. The hypothesis about the role of suprasegmental information in fetal auditory perception has been investigated (Cooper and Aslin, 1989).

In an effort to test whether prenatally available suprasegmental information would be sufficient to induce a postnatal preference, the authors had 13 pregnant women sing the lyrics of the tune to "*Mary Had A Little Lamb*" using the syllable "*la*" instead of the actual words of the melody (Cooper and Aslin, 1989). Each woman sang the melody 5 minutes daily starting on the 14th day prior to her due date. The newborns of these mothers were tested between 34 and 72 hours after birth (mean age = 52 hours old) using the IBI procedure. For the seven infants in the prenatal group, sucking bursts that ended IBIs < mIBI turned on a recording of "*Mary Had A Little Lamb*" sung by a professional female singer (using "*la*" instead of the words), whereas sucking bursts that ended IBIs  $\geq$  mIBI turned on a recording of the same singer singing "*Love Somebody*", also with "*la*" instead of the words. These two melodies were sung in the same key and contained the

same absolute notes, but the notes occurred in different orders to yield different melodic contours. For the other six infants in the prenatal group, the IBI/melody contingency was reversed. In addition, a control group of eight newborns was tested under the identical condition except that they had no prior experience with either melody. The results showed that the newborns in the prenatal group produced more of the IBIs to turn on their familiar melody compared to their baseline performance, while the newborns in the control group did not, regardless of condition. This study demonstrated that the suprasegmental characteristics of a prenatally experienced melody were sufficient to induce a postnatal preference for that melody.

Further supporting evidence for the salience of suprasegmental information in fetal perception comes from the demonstration that newborns discriminated and preferred their native language to a foreign language (Mehler et al., 1988; Moon, Cooper and Fifer, 1993). Using the /a/ or /i/ signal discrimination procedure (Moon and Fifer, 1990), Moon et al. (1993) demonstrated that 2-day-old newborns whose mothers were monolingual speakers of Spanish or English, preferred their mother's language to the other one. Demonstration of a preference for the native language at such an early age favors an interpretation of the study by Mehler et al. (1988) in terms of a prenatal familiarization. In the latter studies, using a noncontingent habituation / dishabituation of high-amplitude sucking procedure, Mehler et al. (1988) demonstrated that 4-day-old native French newborns could discriminate a recording of a woman speaking Russian from the same woman speaking French, but did not differentially respond to English from Italian recordings. Also, 4-day-olds of non-French parents did not respond differentially to either Russian or French recordings. Thus, very young infants seem to require some

experience with a language in order to respond differentially to languages. This interpretation is strengthened by additional data (Mehler et al., 1988) showing that native English 2-month-olds also did not respond differentially to Russian or French, but easily discriminated English from Italian. Thus, it was not merely the young age of the newborns that resulted in their failure to respond differentially to nonnative languages. Prenatal maternal speech is one likely source of native language experience for the newborns.

Finally, Mehler et al. (1988) demonstrated that native French 4-day-old newborns and native English 2-month-olds could still discriminate French from Russian and English from Italian, respectively, even when all of these recordings were low-pass filtered at 400 Hz, which effectively removed most segmental information and maintained their intonational and temporal structures. It is more likely that prenatal auditory experience with the suprasegmental features of maternal speech influences the ability of newborns to discriminate their native language from other nonnative language, although it certainly is possible that newborns rely on both segmental and suprasegmental information when discriminating their native language from a foreign language.

There is now clear evidence that from the earliest days of postnatal life the human infant is actively engaged in processing sounds, particularly those containing acoustic attributes of the infant's native language. The infant's prenatal experience with maternal speech may, in large part, determine the early postnatal perceptual salience of a specific mother's speech and native speech.

## Speech Perception

### Speech Perception in Infancy

There are two characterizations of infants' "initial state" regarding speech perception. One argues that infants enter the world equipped with specialized speech-specific mechanisms evolved for the perception of speech, and that infants are born with a "speech module" to decode the complex and intricate speech signals (Foder, 1983; Mehler and Dupoux, 1994). The other holds that infants begin life without specialized mechanisms dedicated to speech, and that infants' initial responsiveness to speech can be attributed to their more general sensory and cognitive abilities (Aslin, 1987; Kuhl, 1987; Jusczyk, 1996).

In fact, the capacity of newborns to distinguish minimal speech contrasts is remarkable (Aslin, Pisoni and Jusczyk, 1983; Aslin, 1987; Kuhl, 1987; Mehler and Dupoux, 1994). Eimas et al. (1971) were the first to demonstrate that human infants, as young as one month old, can discriminate subtle acoustic properties in a categorical manner that differentiate for English-speaking adults the stop-consonant-vowel syllables /ba/ from /pa/, which are different in voice onset time (VOT). In their study, computer-generated (synthetic) speech differing only VOT was presented in pairs to infants for testing with the high-amplitude sucking procedure. Only one of these VOT pairs spanned the boundary between English-speaking adults' phonemic categories for /ba/ and /pa/. This between-category VOT pair was discriminated by the infants, whereas several other within-category pairs were not discriminated, even though the VOT difference between each pair was identical (20 second). Since then, there is growing body of evidence that nearly all speech contrasts (phonetic contrasts) used in any of the world's natural



languages can be discriminated by 6 months of age (Aslin, Pisoni and Jusczyk, 1983; Aslin, 1987; Kuhl, 1987; Jusczyk, 1996). There are also indications that during the early stages, the mechanisms that underlie speech processing by infants may be a part of more general auditory processing capacities (Aslin, Pisoni and Jusczyk, 1983; Aslin, 1987; Kuhl, 1987; Jusczyk, 1996). Prior to 6 months of age, infants are performing their analysis of speech sounds solely on the basis of acoustic differences. These acoustic differences are sufficient to permit categorical perception, just as similar acoustic mechanisms presumably support the processing of nonspeech contrasts by infants (Jusczyk et al., 1983) and the processing of speech contrasts by nonhumans (Kuhl and Miller, 1975, 1978).

### Characteristic of Speech

Speech signals have numerous distinctive acoustic properties or attributes that are used in the earliest stages of perceptual analysis. The average intensity of normal speech, measured at a distance of 30 centimeter from the speaker's lips, is about 66 dB intensity level (IL), and individual variation between speakers is about  $\pm 5$  dB (Dunn and White, 1940). If the pauses (silent intervals) are excluded, the experimental data indicated that these levels would be increased 3 dB (Fletcher, 1953). Loud speech may reach 86 dB IL, while soft speech may be as low as 46 dB. In the course of ordinary conversation, the dynamic range of speech is about 35-40 dB (Fletcher, 1953). In a more recent study (Cox and Moore, 1988), the mean sound pressure level at 1 meter for a male talker speaking with normal vocal effort was 61 dB and for a female talker was 59 dB. The average spectra were similar in the range from 400 to 5000 Hz between male and female talkers.

Interestingly, the comparison of long-term average speech spectra over 12 languages showed that the spectrum was similar for all languages although there were many small differences (Byrne, et al., 1994). The average value of sound pressure level at 20 centimeter for males was 71.8 dB SPL, while that for females was 71.5 dB SPL. For one-third octave bands of speech, the maximum short-term r.m.s. level was 10 dB above the maximum long-term r.m.s. level, and was consistent across languages and frequency.

Most of the energy of speech derives from vowels. Vowels are usually more intense and relatively longer in duration than consonants. The average difference in intensity between vowels and consonants is about 12 dB. In English, the intensity difference between the weakest consonants /θ/ and the strongest vowel /o/ is about 28 dB (Fletcher, 1953). The frequency range of speech extends from 80 Hz to several thousand Hertz, while the frequencies important to the speech signal are within the 100 to 5000 Hz range (Borden and Harris, 1984). The human voice is composed of many frequencies. The lowest frequency is the fundamental frequency of the voice, driven by the vibration of the vocal folds. The fundamental frequency is constantly changing during articulation, and varies considerably from one person to another. The fundamental frequency of a low-pitched male voice is about 90 Hz, while a woman with a high-pitched voice may speak at a fundamental frequency of about 300 Hz. On average, the average female voice corresponds to middle C or 256 Hz, whereas the male voice is about an octave lower (Fletcher, 1953).

The energy in vowels is concentrated mainly in the harmonic sounds of the fundamental frequency, which for each vowel is divided into several typical frequency regions, called formants, whose center frequency depends on the shape of the vocal tract

(resonance of the vocal tract). In addition to the fundamental frequency ( $F_0$ ), four formants are usually recognized; the lowest two formants ( $F_1$  and  $F_2$ ) are stronger than the other two and occur at frequencies typical for each vowel. The lowest three formants are the most important for correct recognition of English vowels. The frequency range of these formants fits fairly well within the 300-3500 Hz range, which is the standard bandwidth used in the telephone industry (Borden and Harris, 1984; Kent, 1997). If the fundamental frequency is raised by an octave, the formant values increase by only 17 percent (Peterson and Barney, 1952).

The consonants differ essentially from the vowels in that they usually have no distinct formant composition; they are composed of mostly high-frequency noise components. In most consonants, however, energy is concentrated mainly in characteristic frequency regions. Thus, consonant sounds have components that are higher in frequency and lower in intensity than vowel sounds. The intensity tends to be scattered continuously over the frequency region characteristic of each consonant sound (French and Steinberg, 1947; Borden and Harris, 1984; Kent, 1997).

In contrast to acoustic phonetics that identifies speech sounds in terms of acoustic parameters (frequency composition, relative intensity, and duration changes), traditional phonetics describes speech sounds in terms of the way they are produced. The main divisions are voicing, place and manner. "Voicing" is related to vocal fold vibration, e.g., voiced or voiceless. "Place" is related to the location of the major airflow constriction of the vocal tract during articulation, e.g., bilabial, labio-dental, lingui-dental, alveolar, palatal or velar. "Manner" is related to the degree of nasal, oral, or pharyngeal cavity

construction, e.g., vowels, stops (plosives), nasals, fricates, affricates, liquids or glides.

Thus, /b/ in the word "best" is a voiced bilabial stop (plosive) (Borden and Harris, 1984).

### Intelligibility of Speech

The ability to understand speech is the most important measurable aspect of human auditory function. Speech can be detected as a signal as soon as the most intense point of its spectrum exceeds the ear's pure tone threshold at the frequency concerned. This intensity is called the speech detection threshold or threshold of detectability (Egan, 1948; Schill, 1985). At this intensity level, a listener is just able to detect the presence of speech sounds about 50% of the time. When the intensity is increased by some 8 dB, the subjects begin to understand some words and can repeat half of the speech material presented; this is the speech reception threshold or threshold of perceptibility (Egan, 1948; Hawkins and Stevens, 1950; Schill, 1985). The speech reception threshold of spondee words (two syllables), which is considerably lower than one-syllable words, is at about 20 dB SPL (Davis, 1948; Penord, 1985). However, only after the average intensity of speech has reached between 30 to 33 dB SPL, are 50 percent of monosyllabic words understood (Kryter, 1946; French and Steinberg, 1947; Davis, 1948; Egan, 1948). Speech intelligibility or speech discrimination, expressed in terms of percentage correct, is used to describe how much speech sound can be understood. The factors affecting speech intelligibility are numerous. These include physical factors related to the speech stimuli such as level of presentation, frequency composition, distortion, and signal to noise ratio.

French and Steinberg (1947) used nonsense monosyllables of the consonant-vowel-consonant (CVC) type as word material in their studies, and examined intelligibility after low-pass and high-pass filtering. They found that when intensity was increased, discrimination improved up to a certain limit, after which it remained largely constant even if intensity was further increased. Optimal intensity with different filter settings proved to be approximately the same, within a range of 10 dB. The optimal intensity was 75 dB SPL. At this level, when all frequencies above 1000 Hz were passed through the filter, 90% of CVC syllables were recognized correctly. However, when only the frequencies below 1000 Hz were presented, correct identification of the CVC syllables declined to 27%. The French and Steinberg study clearly demonstrated the importance of the high frequencies for correct identification of CVC syllables. Furthermore, when intelligibility scores were plotted as a function of cutoff-frequency of at optimal intensity levels, the low-pass and high-pass curves intersected at 1900 Hz, where the intelligibility score was 68%. It was said that the crossover point divided the frequency scale into two equivalent parts; the frequencies above the cross were as important as the frequencies below the crossover frequency.

The type of speech material distinctly affects the intelligibility of filtered speech (Hirsh, Reynolds and Joseph, 1954). The speech materials in their study included nonsense syllables, monosyllabic words (Central Institute for the Deaf Auditory Test W-22), disyllabic words (spondees, iambs and trochees) and polysyllabic words. The input speech level for all filter conditions was 95 dB SPL. They found that nonsense monosyllables and monosyllable words suffered most in intelligibility during frequency filtering. When the cutoff frequency (high-pass filter) was less than 3200 Hz, the

intelligibility did not decrease significantly. But intelligibility decreased rapidly as the cutoff frequency increased above 3200 Hz. Under low-pass filter conditions, it was only when all the frequencies above 800 Hz were eliminated that the intelligibility decreased noticeably from its maximum, and then it dropped rapidly as the more extreme filter conditions were reached. The functional curves for the different speech materials remained nearly constant under both high-pass and low-pass filtering. The fewer syllables there were in a meaningful word the lower its intelligibility. Nonsense monosyllables were the least intelligible of all. Intelligibility of nonsense syllables and monosyllable words is severely affected by frequency distortion. However, as word length increases, intelligibility is retained. For nonsense syllables, the low-pass and high-pass functional curves intersected at 1700 Hz, where the intelligibility score was 75%. The higher crossover frequency (1900 Hz) with lower intelligibility score (68%) in the French and Steinberg (1947) curves may be due to the high rejection rate of the filters. Hirsh et al. (1954) also studied noise-masking effects on the intelligibility of different types of speech materials. The intelligibility of easy speech material increased more rapidly as a function of signal-to-noise (S/N) ratio than did the intelligibility of more difficult material. At a given S/N ratio, noise levels significantly affect intelligibility. In general, intelligibility at a noise level of 70 dB was higher than that at other noise levels. The results also showed that the intelligibility of polysyllabic, disyllabic and monosyllabic words in noise was higher when they appeared in sentences than when they appeared as discrete items on a list. Differences among the intelligibility of the different types of words were much smaller when the words appeared in sentences. Sentence context had the greatest benefit on understanding monosyllabic words.

Pollack (1948) increased the difficulty of the test method for studying the effect of low-pass and high-pass filtering by adding continuous spectrum white noise at 81.5 dB SPL as a constant background noise. The test material consisted of monosyllabic, phonetically balanced words. The overall speech level was about 68 dB SPL at a distance of 1 meter from the talker. In general, the results indicated that speech intelligibility increased as the intensity level of the speech signal and the frequency range were increased. Owing to the background noise, +10 dB orthotelephonic gain (ratio of the sound intensity at the listener's ear produced by the test system to the orthotelephonic reference system, about 75 dB SPL) gave only 30 percent discrimination even to unfiltered speech. With low-pass and high-pass filtering, the intelligibility improved continuously with increasing intensity, up to a +50 dB orthotelephonic gain with different filter settings, even though the rise of the curves between orthotelephonic gain of +30 and +50 dB was fairly slight. The introduction of background noise resulted in shifting optimal intensity from +10 dB orthotelephonic gain (French and Steinberg, 1947) to the +30 to +50 dB level.

The Pollack (1948) study also demonstrated that the contribution to the intelligibility of the higher speech frequencies alone was small. When a high-pass filter with a 2375 Hz cutoff was used, intelligibility was only 5% at maximal gain. However, these same frequencies made an appreciable difference in intelligibility when the low frequency sounds were also passed at the same time. When the cutoff frequency of low-pass filter was extended from 2500 Hz to 3950 Hz, the intelligibility was improved from 70% to 90%. It was suggested that the contribution to intelligibility of a given band of speech frequencies was not independent of the contribution being made at the same time

by other bands of frequencies. There was an interaction among the contributions of the various bands. Similarly, the contribution to intelligibility of very low speech frequencies was also small. No words were recognized when the frequencies below 425 Hz alone were heard. However, when high-pass cutoff frequency was decreased from 580 Hz to 350 Hz, the intelligibility was improved from 85% to 93%.

A study of the effects of noise and frequency filtering on the perceptual confusions of English consonants revealed that noise and low-pass filtering ensured more homogeneous and well-defined results, whereas the mistakes from high-pass filtering were more indefinite (Miller and Nicely, 1955). Nonsense consonant-vowel (CV) syllables were used as the test material. The 16 consonants were spoken initially before the vowel /a/. The results showed that voicing and nasality (manner of articulation) were much less affected by a random masking noise than were the other features. Affrication and duration (manner of articulation) were somewhat superior to place but far inferior to voicing and nasality. Voicing and nasality were discriminable at S/N ratio as poor as -12 dB whereas the place of articulation was hard to distinguish at S/N ratio less than 6 dB, an 18 dB difference in efficiency. After low-pass filtering (cutoff frequency ranged from 5000 Hz to 300 Hz), voicing and nasality features were well preserved compared with affrication and place information although affrication was superior to place of articulation. These results showed the considerable similarity between masking by broadband noise and filtering by low-pass filters. The authors explained that the uniform noise spectrum masked high frequencies more than low frequencies since the high-frequency components of speech were relative weaker than low-frequency components, so it was in effect a kind of low-pass filter. However, high-pass filtering (cutoff



frequency ranged from 1000 Hz to 4500 Hz) produced a totally different pattern. All features deteriorated in about the same way as the low frequencies were removed. Thus, low-pass filters affected linguistic features differentially, leaving the phonemes audible but similar in predictable ways, whereas high-pass filters removed most of the acoustic power in the consonants, leaving them inaudible and producing quite random confusions. Audibility was the problem for high-pass filtering and confusibility was the problem for low-pass filtering. In addition, the crossover point of the high-pass and low-pass function curves was 1550 Hz, and it became 1250 Hz when plotted by the relative amount of information transmitted instead of the intelligibility score. The downward shift of crossover point in frequency indicated that relative to the intelligibility, the low-pass information was greater and the high-pass information was smaller in consonant recognition.

Wang and her colleagues studied perceptual features of consonant confusions in noise (Wang and Bilger, 1973), and following filtering distortion of speech (Wang, Reed and Bilger, 1978), by sequential information analysis (SINFA), which sequentially identifies features with a high proportion of transmitted information contributing to consonant perception. Nonsense syllables were used as test materials in their studies. The stimuli represented all phonologically permissible consonant-vowel (CV) and vowel-consonant (VC) syllables, which were formed by combining one of 25 consonants with the vowels /i/, /a/ or /u/. Wang and Bilger (1973) demonstrated that articulatory and phonological features could account for a large proportion of transmitted information. The particular features, which resulted in high levels of performance, varied significantly from one syllable set to another and in some cases varied within syllable sets as a

function of listening conditions. Voice and nasal features were well perceived both in noise and in quiet, and they were identified as perceptually important in every syllable set where they were distinctive. The feature round (/w/ and /h<sup>w</sup>/) was also well perceived both in noise and in quiet. Other features, such as frication and place, appeared to have different perceptual importance depending upon the listening condition. Under filtering conditions, there were differential effects of high-pass and low-pass filtering on feature recognition (Wang, Reed and Bilger, 1978). Low-pass filtering (cutoff frequency ranged from 5600 Hz to 500 Hz) produced systematic changes in the importance of different features, whereas high-pass filtering (cutoff frequency ranged from 355 Hz to 4000 Hz) produced less consistent changes in features recognition. When the low-pass cutoff was lowered from 2800 to 1400 Hz, sibilance (/s/, /z/, /S/, /tS/, /Z/ and /dZ/) (manner of articulation) quickly lost its perceptibility. The high-pass filtering had little effect on the recognition of sibilance. The high crossover point of the functions at 2800 Hz indicated that cues for sibilant sound lay in the high-frequency region of the spectrum, above 2000 Hz. High (/k/, /g/, /S/, /tS/, /Z/, /dZ/, /ŋ/, /w/ and /j/) and anterior (/p/, /t/, /b/, /d/, /t/, /s/, /v/, /z/, /m/, /n/, /l/, /θ/ and /ð/) features (place of articulation) also dropped noticeably when the cutoff of low-pass filter was lowered to 1400 Hz. For CV syllables, the crossover point, approximately 1700 Hz, was lower than that for VC syllables, about 2400 Hz. Thus, the cues for high / anterior features were partly dependent on the position of the consonant within the syllables. However, voice and nasality became increasingly important as the low-pass cutoff was lowered, while they were adversely affected by high-pass filtering. The characteristics of consonant confusions following filtering were quite similar to that noted by Miller and Nicely (1955).

The patterns of consonant confusions generated by subjects with sensorineural hearing loss were like those generated by normal hearing subjects in response to the appropriate filtering distortion of speech (Bilger and Wang, 1976; Wang, Reed and Bilger, 1978). For example, severe low-pass filtering produced consonant confusions comparable to those of listeners with high-frequency hearing loss. Severe high-pass filtering gave a result comparable to that of patients with flat or rising hearing loss.

In 1994, Griffiths et al. investigated the intelligibility of speech stimuli recorded within the uterus of a pregnant sheep. The results showed that the intelligibility of the phonemes recorded in the air was significantly greater than the intelligibility of phonemes recorded *in utero*. A male talker's voice was more intelligible than a female talker's voice when the recordings were made *in utero*. Furthermore, an analysis of the feature information transmission from recordings inside and outside the uterus revealed that voicing information is better transmitted *in utero* than place or manner information. The findings are quite similar to those of studies conducted by Miller and Nicely (1955) and Wang et al. (1978) in that transmission into the uterus can be modeled as a low-pass filter. While the results of Griffiths et al. (1994) study only reflect the perceptibility of the speech energies present in the amniotic fluid, they do not specify what speech energy might be present at the level of fetal inner ear. Measurements of acoustic transmission to the fetal inner ear are quite limited at present. The purpose of current study was to evaluate the intelligibility of externally generated speech utterances transmitted to and recorded at the fetal sheep inner ear *in utero*.

### CHAPTER 3 MATERIALS AND METHODS

The overall aims of this project were to determine the intelligibility of speech information that was transmitted into the uterus and present within the inner ear of the sheep fetus *in utero*. Cues inherent in the speech of both the mother and external talkers may be perceived by the fetus, thus forming the basis for language acquisition. This study was intended to provide evidence of fetal inner ear physiological responses to externally generated speech and to address the hypotheses included in Chapter 1. The study had two distinct components. The first involved recording speech produced through a loudspeaker with an air microphone, a hydrophone placed in the uterus of a pregnant sheep and an electrode secured to the round window of the fetus *in utero* (cochlear microphonic, CM). The second portion of the study involved playing the recordings to a jury of normal hearing adults so speech intelligibility could be evaluated.

#### Surgery

Eight time-mated pregnant ewes carrying fetuses at gestational ages from 130-140 days were prepared for surgery (term is 145 days). From this group, speech stimuli recorded from only one animal were used in this study. Recordings from this animal were judged by the experimenter to have the best fidelity. Speech signals produced from a

loudspeaker were recorded with an air microphone, a hydrophone placed in the uterus of pregnant sheep and an electrode secured to the round window of the fetus. The Animal Use Protocol in this study was approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Florida.

In preparation for measurements of fetal cochlear microphonic (CM), ewes were fasted, anesthetized and maintained on a mixture of oxygen and halothane (1.5-2%) during surgery and subsequent experimentation. The ewe was placed in the supine position and the fetal head was delivered through a midline hysterotomy. An incision was made over the fetal right bulla posterior and inferior to the pinna. The incision was located at the attachment of the cartilaginous portion of the canal to the lateral surface of the skull and was made parallel to the posterior border of the mandibular ramus. The bulla was exposed and a small hole was opened through the bulla. The round window was located with an operating microscope. An electrode was made from insulated stranded stainless steel wire (Cooner Wire Company, Chatsworth, CA) with the insulation removed from one end. The uninsulated end was rolled into a 2-mm diameter ball and placed inside the round window niche (positive electrode). After verifying the impedance of the round window electrode ( $< 10 \text{ k}\Omega$ ), the bulla was refilled with amniotic fluid and sealed over with methylmethacrylate. Additional Cooner wire electrodes were sutured to tissue overlying the bulla (negative electrode) and to tissue at a remote site (ground electrode). The skin over the bulla was sutured and the electrodes were carefully secured to the fetus with silk thread. The fetus was returned to the uterus and the uterus and abdomen were closed with clamps. Electrode

wires passed through the incisions and were connected to a biological amplifier (Grass Instruments Co., model P511K, Quincy, MA).

### Recording Speech Stimuli

The anesthetized ewe was placed supine on a stretcher and transported to a sound-treated booth (Industrial Acoustics Co., model GDC-1L, Bronx, NY). Speech stimuli for producing fetal CM were prerecorded on cassette tape and consisted of Vowel-Consonant-Vowel (VCV) nonsense syllables and Consonant-Vowel-Consonant (CVC) monosyllable words spoken by a male and a female talker. The center of a loudspeaker was one meter from the ewe and was adjusted to the same height as the center of the lateral wall of the ewe's abdomen. A calibrated air microphone (Brüel and Kjael, type 4165, Marlborough, MA) was positioned over the maternal abdomen at a distance of 10 cm. A miniature hydrophone (Brüel and Kjael, model 8103), calibrated with a pistonphone (Brüel and Kjael, model 4223), was inserted in the uterus and connected to a charge amplifier (Brüel and Kjael, type 2635). The output from the tape player (Harman Kardon, model TD 392, Woodbury, NY) was routed through a power amplifier (Peavey DECA/1200, Peavey Electronics Corp., Meridian, MS) that activated the loudspeaker (Peavey HDH-2). The cochlear potentials, CMs recorded from the fetal inner ear in response to the speech stimuli, were amplified (Grass Instruments Co., model P511K, Quincy, MA) and high-pass filtered at 100 Hz (Kron-Hite Corp., model 3550, Avon, MA, 24 dB/octave). Figure 3-1 showed the schematic drawing of recording system set-up.

Because the CM is produced during acoustic stimulation, the potential can be contaminated with electromagnetic artifact emanating from the loudspeaker and associated

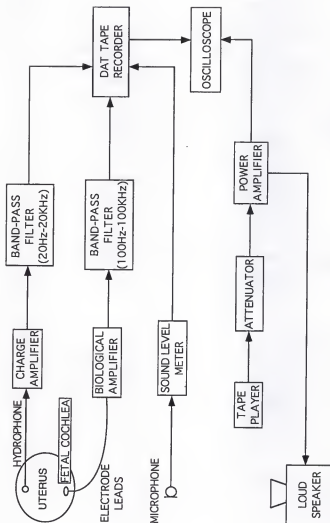


Figure 3-1. Schematic drawing showing the animal and the setup of devices for stimulus generation, stimulus measurement, and recording in air, in the uterus, and from the fetal inner ear (cochlear microphonic).

wires. The electrical interference produces a voltage output from the biological amplifier that mimics the true biologic potential. Because electromagnetic energy travels at the speed of light, whereas acoustic energy travels at the speed of sound (344 m/s), uncontaminated CM occurred approximately 3 ms after the onset of the stimulus. If this onset delay was not present in the recording, then measurements were repeated after appropriate equipment adjustment and / or grounding. The presence of an onset delay confirmed that the recorded waveform was bioelectric rather than electromagnetic (Gerhardt et al., 1992).

Before recording speech stimuli, CMs (Figure 3-2) were verified by using tone-bursts (0.5, 1.0 and 2.0 kHz). An evoked potential averaging computer (Tucker-Davis Technologies, Gainesville, FL) delivered stimuli to the loudspeaker. Tone bursts were delivered to the ewe's flank at intensity levels that were capable of producing CM responses. Twenty stimuli were delivered and averaged for each CM response. Stimulus duration (10 or 20 ms), sweep time (20 or 50 ms) and filtering (100-3,000 Hz or 100-10,000 Hz) varied with stimulus frequency (0.5, 1.0 and 2.0 kHz). The rate of stimulation was 5/s and the rise/fall time was 0.2 ms.

The speech stimuli were delivered to the flank of pregnant ewes at two intensity levels (105 and 95 dB SPL). First, the signals were simultaneously detected with a microphone located over the abdomen and electrodes placed on the fetal round window *in utero*. The outputs from the microphone and inner ear (CM) were recorded on two separate channels of a DAT tape recorder (SONY Corporation, type Z45ES, Japan). Then, the same speech stimuli were repeated and recorded with a hydrophone placed in the uterus and electrodes placed on the fetal round window *ex utero*. The fetal external canal and middle ear cavity were cleared of fluids during *ex utero* measurement. At the completion of all



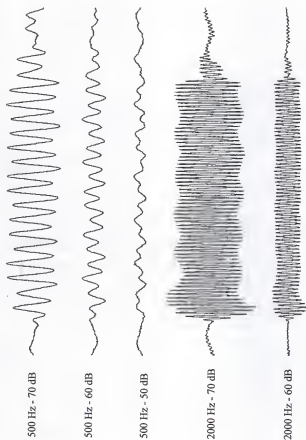


Figure 3-2. CM responses obtained from a fetal sheep. Examples of CMs evoked by airborne pure tones at 0.5 and 2.0 kHz and at stimulus levels indicated under each waveform. The apparent onset latency represents the acoustic travel-time from the loudspeaker to the fetal inner.

measurements, the ewe and fetus were euthanized as prescribed by the IACUC of the University of Florida.

### Perceptual Testing

#### Subjects

A total of 155 undergraduate students from the Department of Communication Sciences and Disorders at University of Florida volunteered to participate in this study. From this group, responses from 139 students who judged the intelligibility of speech stimuli were used. Sixteen students were excluded from the study for the following reasons: eight judges used unreadable symbols; four judges were nonnative American English speakers; and four judges reported hearing loss. The descriptive information of the perceptual tests is presented in Table 3-1.

All of the judges had taken or were taking an undergraduate course in phonetics, although as a group they would not be considered experienced phoneticians. All testing was completed in a single 45-minute session. The protocol for the perceptual testing was approved by the University of Florida Institutional Review Board (UFIRB Project # 1998-563).

#### Speech Stimuli

Two sets of stimuli were used, vowel-consonant-vowel (VCV) nonsense syllables and consonant-vowel-consonant (CVC) words spoken by male and female talkers and words based on the Griffiths word lists (1967). Each stimulus item was presented in a

Table 3-1. Perceptual tests.

| Perceptual audio CD | Contents | Number of judges |
|---------------------|----------|------------------|
| A                   | VCV      | 33               |
| B                   | CVC      | 19               |
| C                   | CVC      | 21               |
| D                   | CVC      | 20               |
| E                   | CVC      | 21               |
| F                   | CVC      | 25               |

carrier phrase, "Mark the word \_\_\_\_." The 14 nonsense syllables (C=/p, t, k, b, d, g, f, v, s, z, m, n, S, tS/) spoken by both a male and a female talker were preceded and followed by the vowel /a/ (e.g. /aga/). The mean fundamental frequencies were 120 and 225 Hz for the male and female talkers, respectively. Sixty-four items were recorded at each of 16 conditions among gender of talker (male and female), stimulus levels (105 and 95 dB SPL), and recording locations (air, uterus, CM *ex utero*, and CM *in utero*).

### Procedures

The word list, spoken by both male and female talkers, were played through the loudspeaker via a cassette tape recorder at two different airborne levels measured at the maternal flank: 105 and 95 dB SPL (dB *re*: 20  $\mu$ Pa). The outputs from the air microphone, the hydrophone, and the fetus inner ear (CM) *ex utero* and *in utero* were recorded on DAT tapes. One set of recordings with the best quality sound from one fetus was chosen for constructing perceptual tapes. First, speech stimuli were digitized and reproduced via a computer program (Cool Edit, Syntrillium Software Corporation, Phoenix, AZ) with 44.1-kHz sampling rate and 16-bit resolution. The amplitudes of the speech stimuli were adjusted to the same relative voltage levels. Second, each syllable item with a carrier phrase was saved as an individual file. Then a computer program was used to randomize and counter-balance the speech stimuli among gender of talker (male and female), stimulus levels (105 and 95 dB SPL), and recording locations (air, uterus, CM *ex utero*, and CM *in utero*). Finally, six different perceptual audio compact discs (CDs) were created. One contained randomized recordings of 224 nonsense items (14 nonsense syllables recorded

under 16 conditions). The five other CDs contained recordings of 800 monosyllabic words, each version consisted of 160 words (10 words recorded under 16 conditions, the same word occurred no more than 4 times in each version). A 5-second silence interval separated each test item.

The recordings were used to conduct a perceptual test of speech intelligibility. The test required groups of judges to listen to the utterances in the carrier phrase and mark on paper what they heard. The judges' responses provided the basis for determining intelligibility scores (percent correct) associated with the VCV nonsense items and the CVC words.

For the 14 VCV nonsense items, the judges filled in a blank in a /a\_\_a/ frame with the vowel set to /a/. For example, if a judge heard "Mark the word /apa/," he or she would have to write a "p" in the blank to be correct.

For the 50 CVC words, each judge selected his or her response from a closed set of six monosyllabic words that differed in either the initial or final consonant. For example, one stimulus item was "Mark the word bat" and the response list included "batch, bash, bat, bass, back, badge." To be correct, the judge would have to mark the word "bat."

Each version of perceptual audio CDs were played to a group of judges comprising 20-30 normal hearing young adults. All testing were conducted in a specially designed listening laboratory which accommodated up to 25 people at one time. The perceptual audio CD were played over earphones (HS-95 and HS-56, SONY) to the judges at an output level set to be comfortably loud (approximately 70 dB SPL). Figure 3-3 showed the frequency responses of two types of earphones used in the perceptual tests. Each listening

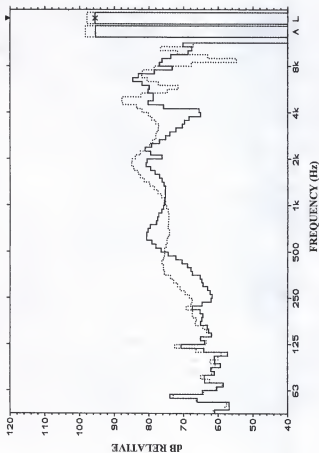


Figure 3-3. The frequency responses of two types of earphones: SONY HS-95 (dot line) and HS-56 (solid line) used for the perceptual tests.

test was preceded by a brief practice session using a version of perceptual audio CD different from the real testing CD to ensure that subjects understood the perceptual tests.

### Data Analyses

#### Statistical Analyses

Intelligibility, consonant confusion matrices and spectral analyses of recorded speech signals were assessed. The speech intelligibility scores (percent correct) were derived from the judges' responses to the perceptual audio CDs for the VCV nonsense syllables and CVC words by gender, intensity level, and recording location. Multifactor analysis of variance (ANOVA) was performed on the data of the VCV nonsense syllables and CVC words separately. The independent variables included three factors: gender of the talker (male and female), sound pressure level of the airborne stimulus (105 and 95 dB), and location of recording (air, uterus, CM from *ex utero* fetus, and CM from *in utero* fetus). The dependent variables were percentage of correct identification of nonsense syllables and monosyllabic words (perceptual scores). In order to meet the variance assumptions for statistical analysis, the percent intelligibility data, which are binomial variables (Thornton and Raffin, 1978), were transformed using an arcsine function ( $2 \times \arcsin \sqrt{(\%/100)}$ ) to normalize the variance prior to further analysis (Winer, Brown and Michels, 1991).

### Information Analyses

Data were presented in the form of a  $14 \times 14$ -item confusion matrix for each condition. A total of 16 matrices for VCV nonsense syllables were collected. Sequential Information Analysis (SINFA; Wang, 1976) of perceptual pattern was performed. SINFA is applied to the error matrices in order to evaluate the amount of feature information received. SINFA allows for the partitioning of the contingent information transmitted and received for particular features of the stimuli (e.g., voicing, manner, and place). From these results a relative measure of performance may be calculated (the ratio of the bits of information received to the bits sent, with the effects of other features held constant). The data from all 16 conditions were analyzed using SINFA.

### Acoustic Analyses

Acoustic analyses of five vowels (/i/, /I/, /e/, /æ/, /A/) selected from the Griffiths' words list (CVC) were performed across the recording conditions (105-dB stimuli of both male and female speakers recorded in air, in the uterus, CM from *ex utero* fetus, and CM from *in utero* fetus). The fundamental frequency ( $F_0$ ) and the first three formant frequencies ( $F_1$ ,  $F_2$ , and  $F_3$ ), and their relative intensity levels were measured by using a signal-processing computer program (Cool Edit, Syntrillium Software Corporation, Phoenix, AZ). Each real-time speech waveform was digitized with 44.1-kHz sampling rate and 16-bit resolution. An average 150-ms segment was selected around the steady-state portion of each vowel. The  $F_0$  and formants ( $F_1$ ,  $F_2$ , and  $F_3$ ) of each segment were measured by visual inspection of the corresponding Fourier transform spectrum using Hamming window with



4096 Fourier size followed by smoothing (Lee, Potamianos and Narayanan, 1999).

According to the values measured by Peterson and Barney (1952), and Hillenbrand et al. (1995),  $F_0$  and formants frequencies ( $F_1$ ,  $F_2$ , and  $F_3$ ) were estimated. The relative intensity levels were also calculated by subtracting the background noise value from the peak value under different recording conditions. Two-factor repeated measures ANOVAs were performed on the data of relative intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  across the recording locations for each vowel.

## CHAPTER 4 RESULTS AND DISCUSSION

One hundred and thirty-nine judges completed the perceptual tests. Because the speech stimuli were completed randomized and counter-balanced across gender of talkers (male and female), stimulus levels (105 and 95 dB SPL) and recording locations (air, uterus, CM *ex utero*, and CM *in utero*), learning effects were minimized.

### Intelligibility

The speech intelligibility scores (percent correct) derived from the judges' responses to the perceptual audio compact discs (CDs) for the VCV nonsense syllables and CVC words are displayed in Figures 4-1 and 4-2, respectively. A few general observations can be made about both Figures. First, intelligibility scores as a function of location alone, decreased from air to hydrophone locations and decreased again from CM *ex utero* to CM *in utero*. That is to say, intelligibility scores of the VCV and CVC lists were high when recorded in air and slightly less when recorded with a hydrophone in the uterus. The scores, when recorded from the inner ear of the fetus *ex utero*, are 20-40% lower than recordings from either the air or hydrophone locations. The intelligibility scores recorded from the inner ear of the fetus *in utero* are about 10-20% poorer than the scores recorded from the fetal CM *ex utero*. Second, from casual inspection of the two Figures, there appear to be a slight gender and level effects primarily for the VCV lists.

Figure 4-1. Mean percent intelligibility of VCV nonsense stimuli spoken by a male and a female talker recorded in air, in the uterus, from the fetal CM *ex utero*, and from fetal CM *in utero* at two airborne stimulus levels. Bars equal the standard error of the mean.

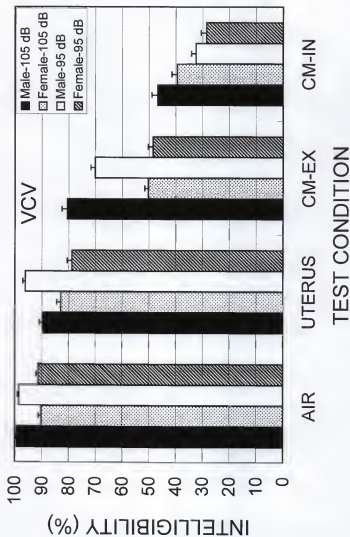
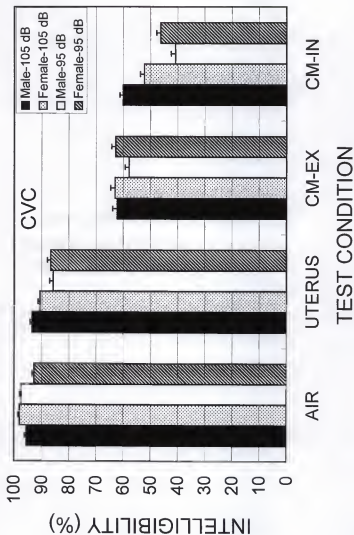


Figure 4-2. Mean percent intelligibility of CVC words spoken by a male and a female talker recorded in air, in the uterus, from the fetal CM *ex utero*, and from fetal CM *in utero* at two airborne stimulus levels. Bars equal the standard error of the mean.



Gender and level effects are more pronounced from recordings of the CM than from recordings in air or in the uterus. Summaries of the means and standard deviations for intelligibility by gender, stimulus level, and location that contributed to these figures are presented in Tables 4-1 and 4-2.

The results of a three-factor repeated measure ANOVA are summarized for VCV stimuli and given in Tables 4-3. There was a significant three-way interaction among gender, stimulus level, and location ( $F_{3,96} = 14.582, p < 0.0001$ ). The main effects were significant for each of the three factors: location ( $F_{3,96} = 994.982, p < 0.0001$ ), gender ( $F_{1,32} = 210.258, p < 0.0001$ ), and stimulus level ( $F_{1,32} = 25.869, p < 0.0001$ ). The results of the post hoc multiple comparison test (Newman-Keuls) are presented in Table 4-4. Not all of the paired results were included in this table. Note that intelligibility in all cases was significantly greater ( $p < 0.01$ ) for CM *ex utero* than for CM *in utero*. Also, intelligibility of the nonsense syllables (VCV) was better at higher presentation levels than at lower presentation levels. When both stimulus levels were compared, statistical significance ( $p < 0.01$ ) was attained for the male voice recorded in the uterus, from CM *ex utero*, and from CM *in utero*, as well as for the female voice recorded from CM *in utero*.

The ANOVA results for CVC words (Table 4-5) showed a significant three-way interaction among gender, stimulus level, and location ( $F_{3,315} = 22.459, p < 0.0001$ ). This was similar to the results for the nonsense syllables (VCV). The main effects were significant for location ( $F_{3,315} = 1213.579, p < 0.0001$ ) and stimulus level ( $F_{1,105} = 102.82, p < 0.0001$ ), but not for gender ( $F_{1,105} = 1.247, p = 0.267$ ). The results of the post hoc multiple comparison test (Newman-Keuls) are given in Table 4-6, in which not all of

Table 4-1. VCV stimulus intelligibility scores for each talker, stimulus level and recording site.

| Male talker        | In Air |        |  | In Uterus |        |  | CM-ex utero |        |  | CM-in utero |        |  |
|--------------------|--------|--------|--|-----------|--------|--|-------------|--------|--|-------------|--------|--|
|                    | 105 dB | 95 dB  |  | 105 dB    | 95 dB  |  | 105 dB      | 95 dB  |  | 105 dB      | 95 dB  |  |
| Mean (%)           | 99.35% | 98.48% |  | 89.61%    | 96.10% |  | 80.52%      | 70.13% |  | 46.75%      | 32.47% |  |
| S.D. (%)           | 2.09%  | 2.97%  |  | 7.16%     | 5.08%  |  | 11.33%      | 8.83%  |  | 12.25%      | 9.46%  |  |
| No. correct (N=14) | 13.909 | 13.788 |  | 12.545    | 13.455 |  | 11.273      | 9.818  |  | 6.545       | 4.545  |  |
| S.D.               | 0.292  | 0.415  |  | 1.003     | 0.711  |  | 1.587       | 1.236  |  | 1.716       | 1.325  |  |
| No. of judges      | 33     | 33     |  | 33        | 33     |  | 33          | 33     |  | 33          | 33     |  |
| Female talker      |        |        |  |           |        |  |             |        |  |             |        |  |
| Mean (%)           | 90.26% | 91.34% |  | 82.90%    | 78.79% |  | 50.22%      | 48.48% |  | 39.39%      | 28.57% |  |
| S.D. (%)           | 5.87%  | 3.89%  |  | 8.74%     | 10.18% |  | 7.68%       | 10.22% |  | 11.73%      | 11.01% |  |
| No. correct (N=14) | 12.636 | 12.788 |  | 11.606    | 11.030 |  | 7.030       | 6.788  |  | 5.515       | 4.000  |  |
| S.D.               | 0.822  | 0.545  |  | 1.223     | 1.425  |  | 1.075       | 1.431  |  | 1.642       | 1.541  |  |
| No. of judges      | 33     | 33     |  | 33        | 33     |  | 33          | 33     |  | 33          | 33     |  |





Table 4-3. ANOVA summary table for VCV stimuli.

| Source  | Sum of Squares | df | Mean Squares | F       | p-value |
|---|----------------|----|--------------|---------|---------|
| Location  | 180.991        | 3  | 60.330       | 994.982 | <0.0001 |
| Error (Location)                                | 5.821          | 96 | 0.06063      |         |         |
| Gender  | 22.470         | 1  | 22.470       | 210.258 | <0.0001 |
| Error (Gender)                                  | 3.420          | 32 | 0.107        |         |         |
| Level   | 0.894          | 1  | 0.894        | 25.869  | <0.0001 |
| Error (Level)                                   | 1.106          | 32 | 0.03456      |         |         |
| Location $\times$ Gender                        | 3.948          | 3  | 1.316        | 21.539  | <0.0001 |
| Error (Location $\times$ Gender)                | 5.866          | 96 | 0.0611       |         |         |
| Location $\times$ Level                         | 2.738          | 3  | 0.913        | 23.181  | <0.0001 |
| Error (Location $\times$ Level)                 | 3.779          | 96 | 0.03936      |         |         |
| Gender $\times$ Level                           | 0.00407        | 1  | 0.0407       | 0.104   | 0.749   |
| Error (Gender $\times$ Level)                   | 1.249          | 32 | 0.03904      |         |         |
| Location $\times$ Gender $\times$ Level         | 2.180          | 3  | 0.727        | 14.582  | <0.0001 |
| Error (Location $\times$ Gender $\times$ Level) | 4.784          | 96 | 0.04983      |         |         |

Table 4-4. *Post hoc* multiple comparisons (Newman-Keuls test) for VCV stimuli.

|     | Conditions |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     | AMH        | AML | UMH | UML | XMH | XML | IMH | IML | AFH | AFL | UFL | XFL | IFH | IFL |
| AMH |            |     |     |     |     |     |     |     |     |     |     |     |     |     |
| AML | --         |     |     |     |     |     |     |     |     |     |     |     |     |     |
| UMH | **         |     |     |     |     |     |     |     |     |     |     |     |     |     |
| UML | *          | *   | **  |     |     |     |     |     |     |     |     |     |     |     |
| XMH | **         |     | **  | **  |     |     |     |     |     |     |     |     |     |     |
| XML | **         | **  | **  | **  | **  |     |     |     |     |     |     |     |     |     |
| IMH | **         |     | **  | **  | **  | **  |     |     |     |     |     |     |     |     |
| IML | **         | **  | **  | **  | **  | **  | **  |     |     |     |     |     |     |     |
| AFH | **         |     |     |     |     |     |     |     |     |     |     |     |     |     |
| AFL | **         | **  |     |     |     |     |     |     | --  |     |     |     |     |     |
| UFL |            |     | **  | **  |     |     |     | **  | **  | **  | --  |     |     |     |
| XFL |            |     |     |     | **  |     |     | **  | **  | **  | **  | --  |     |     |
| IFH |            |     |     |     |     | **  | **  | **  | **  | **  | **  | **  | **  |     |
| IFL |            |     |     |     |     |     | --  | --  | **  | **  | **  | **  | **  | **  |

Note: A = In Air; U = In Uterus; X = CM-ex utero; I = CM-in utero; M = Male; F = Female; H = 105 dB; L = 95 dB.

--  $p > 0.05$ ; \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

Table 4-5. ANOVA summary table for CVC stimuli.

| Source  | Sum of Squares | df  | Mean Squares | F        | p-value |
|---|----------------|-----|--------------|----------|---------|
| Location  | 505.738        | 3   | 168.579      | 1213.687 | <0.0001 |
| Error (Location)                                | 43.753         | 315 | 0.139        |          |         |
| Gender  | 0.192          | 1   | 0.192        | 1.247    | 0.267   |
| Error (Gender)                                  | 16.154         | 105 | 0.154        |          |         |
| Level   | 9.484          | 1   | 9.484        | 102.820  | <0.0001 |
| Error (Level)                                   | 9.685          | 105 | 0.09224      |          |         |
| Location $\times$ Gender                        | 1.2995         | 3   | 0.433        | 3.456    | 0.0658  |
| Error (Location $\times$ Gender)                | 39.486         | 315 | 0.125        |          |         |
| Location $\times$ Level                         | 2.821          | 3   | 0.940        | 8.857    | <0.0001 |
| Error (Location $\times$ Level)                 | 33.439         | 315 | 0.106        |          |         |
| Gender $\times$ Level                           | 0.119          | 1   | 0.119        | 0.566    | 0.454   |
| Error (Gender $\times$ Level)                   | 22.126         | 105 | 0.211        |          |         |
| Location $\times$ Gender $\times$ Level         | 7.713          | 3   | 2.571        | 22.459   | <0.0001 |
| Error (Location $\times$ Gender $\times$ Level) | 36.061         | 315 | 0.114        |          |         |

Table 4-6. *Post hoc* multiple comparisons (Newman-Keuls test) for CVC stimuli.

|     | Conditions |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     | AMH        | AML | UMH | UML | XMH | XML | IMH | IML | AFH | AFL | UFH | UFL | XFH | IFH |
| AMH |            |     |     |     |     |     |     |     |     |     |     |     |     |     |
| AML | *          |     |     |     |     |     |     |     |     |     |     |     |     |     |
| UMH | *          |     |     |     |     |     |     |     |     |     |     |     |     |     |
| UML |            | **  |     |     |     |     |     |     |     |     |     |     |     |     |
| XMH |            |     | **  |     |     |     |     |     |     |     |     |     |     |     |
| XML |            | **  |     | **  |     |     |     |     |     |     |     |     |     |     |
| IMH |            |     |     |     |     |     |     |     |     |     |     |     |     |     |
| IML |            | **  |     | **  |     | **  | **  |     |     |     |     |     |     |     |
| AFH |            |     |     |     |     |     |     |     |     |     |     |     |     |     |
| AFL |            | **  |     |     |     |     |     |     | **  |     |     |     |     |     |
| UFH |            |     | **  |     |     |     |     |     | **  |     |     |     |     |     |
| UFL |            |     |     | --  |     |     |     |     | **  | **  | --  |     |     |     |
| XFH |            |     |     |     |     |     |     |     | **  | **  | **  |     |     |     |
| XFL |            |     |     |     |     |     |     |     | **  | **  | **  | **  | --  |     |
| IFH |            |     |     |     |     |     | **  | **  | **  | **  | **  | **  | **  | **  |
| IFL |            |     |     |     |     |     | **  | **  | **  | **  | **  | **  | **  | **  |

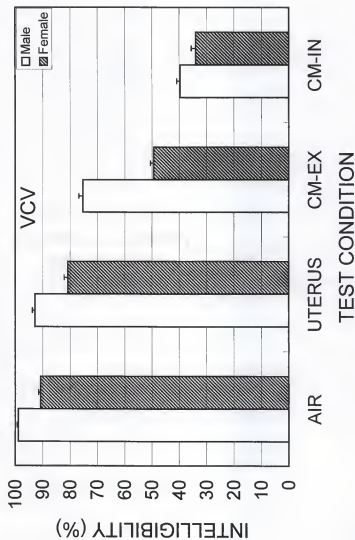
Note: A = In Air; U = In Uterus; X = CM-ex utero; I = CM-in utero; M = Male; F = Female; H = 105 dB; L = 95 dB.

--  $p > 0.05$ ; \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

the paired results were included. It is noted that intelligibility was significantly greater ( $p < 0.01$ ) for CM *ex utero* than for CM *in utero*, except for the male voice recorded at 105 dB SPL ( $p > 0.05$ ). Also, intelligibility of the words (CVC) was better at higher presentation levels than at lower presentation levels, except for the male voice recorded in air. When both stimulus levels were compared, statistical significance ( $p < 0.01$ ) was achieved for the male voice recorded in air ( $p < 0.05$ ), in the uterus, and from CM *in utero*, as well as for the female voice recorded in air and from CM *in utero*.

Figures 4-3 simplifies those data presented in Figure 4-1 by combining levels. For VCV stimuli, the average intelligibility scores for the male voice recorded in air, in the uterus, from fetal CM *ex utero*, and from fetal CM *in utero* were 98.9%, 92.9%, 75.3%, and 39.6%, respectively. For the female voice recorded in air, in the uterus, from fetal CM *ex utero*, and from fetal CM *in utero*, the intelligibility scores were 90.8%, 80.8%, 49.4%, and 34.0%, respectively. A two-factor repeated measures ANOVA indicated significant interaction between gender and location ( $F_{3,96} = 20.925$ ,  $p < 0.0001$ ), and main effects for gender ( $F_{1,32} = 192.744$ ,  $p < 0.0001$ ) and location ( $F_{3,96} = 1048.477$ ,  $p < 0.0001$ ). The post hoc multiple comparison test (Newman-Keuls) indicated that the intelligibility scores of the male voice were significantly higher ( $p < 0.01$ ) than that of the female voice at all four recording locations. Also, for both male and female talkers, the intelligibility scores recorded in air were significantly higher ( $p < 0.01$ ) than that of each of the other three recording locations. The scores recorded in the uterus were significantly higher ( $p < 0.01$ ) than that of recordings from CM *ex utero* and CM *in utero*. The scores recorded from CM *ex utero* were significantly higher ( $p < 0.01$ ) than that from CM *in utero*.

Figure 4-3. Mean percent intelligibility of VCV nonsense stimuli spoken by a male and a female talker recorded in air, in the uterus, from the fetal CM *ex utero*, and from fetal CM *in utero* when combining two airborne stimulus levels. Bars equal the standard error of the mean.

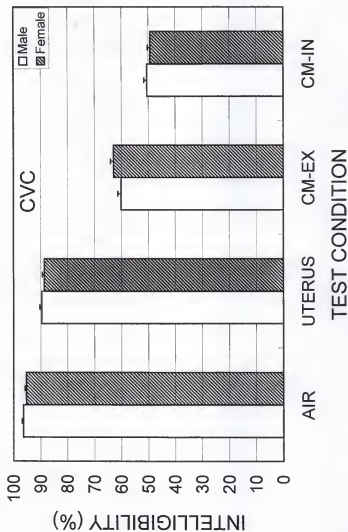




Similarly, Figures 4-4 clarifies those data presented in Figure 4-2 by combining levels. For CVC words, the average intelligibility scores for the male voice recorded in air, in the uterus, from fetal CM *ex utero*, and from fetal CM *in utero* were 96.5%, 89.6%, 60.1%, and 50.5%, respectively. For the female voice recorded in air, in the uterus, from fetal CM *ex utero*, and from fetal CM *in utero*, the intelligibility scores were 95.4%, 88.6%, 62.9%, and 49.3%, respectively. A two-factor repeated measures ANOVA indicated significant interaction between gender and location ( $F_{3, 315} = 3.386, p = 0.0184$ ), and main effects for location ( $F_{3, 315} = 1045.347, p < 0.0001$ ), but not for gender ( $F_{1, 105} = 1.427, p = 0.235$ ). A post hoc multiple comparison test (Newman-Keuls) indicated that, for both male and female talkers, the intelligibility scores recorded in air were significantly higher ( $p < 0.01$ ) than that of each of the other three recording locations. The scores recorded in the uterus were significantly higher ( $p < 0.01$ ) than that of recordings from CM *ex utero* and CM *in utero*. The scores recorded from CM *ex utero* were significantly higher ( $p < 0.01$ ) than that from CM *in utero*. There were no statistical differences ( $p > 0.05$ ) between the male voice and the female voice across recording locations, except when recorded in air ( $p < 0.05$ ).

As reported above, speech (VCV and CVC stimuli) intelligibility scores were significantly higher for the recordings in air than in the uterus. Likewise, the intelligibility was significantly greater for the recordings from CM *ex utero* than from CM *in utero*. The recordings within the uterus reflect the speech energies present in amniotic fluid, whereas the recordings from CM *in utero* represent the actual fetal physiological responses to externally generated speech. The characteristics of transmission of external sound pressure into the maternal abdomen and uterus has been

Figure 4-4. Mean percent intelligibility of CVC words spoken by a male and a female talker recorded in air, in the uterus, from the fetal CM *ex utero*, and from fetal CM *in utero* when combining two airborne stimulus levels. Bars equal the standard error of the mean.



well described in humans (Querleu et al., 1988a; Richards et al., 1992) and sheep (Armitage, Baldwin and Vince, 1980; Vince et al., 1982, 1985; Gerhardt, Abrams and Oliver, 1990). The abdomen wall, uterus, and amniotic fluids can be characterized as a low-pass filter with a high-frequency cutoff at 250 Hz and a rejection rate of approximately 6 dB per octave. For frequencies below 250 Hz, sound pressures passing through to the fetus are unattenuated, and, in some cases, are enhanced. Above 250 Hz, sound pressures are increasingly attenuated by up to 20 dB (Gerhardt, Abrams and Oliver, 1990). Thus, the speech signals would be altered as they passed through tissues of the ewe into the uterus. Additionally, the spectral contents of external sounds are further modified by the route of sound transmission into the fetal inner ear. Sound pressures pass through the fetal head by a bone conduction pathway (Gerhardt et al., 1996). For 125 to 250 Hz, an airborne signal would be reduced by 10-20 dB before reaching the fetal inner ear. For 500 through 2000 Hz, the signal would be reduced by 35-45 dB (Gerhardt et al., 1992). Therefore, the recordings of speech from CM *in utero* would be further degraded and less intelligible than the recordings in air and in the uterus.

The present findings reveal better intelligibility for speech in the uterus than has been previously found (Querleu et al., 1988b; Griffiths et al., 1994). Querleu et al. (1988b) found that about 30% of 3120 French phonemes recorded within the uterus of pregnant women were recognized. In 1994, Griffiths et al. evaluated the intelligibility of speech stimuli (VCV nonsense syllables and CVC words) recorded within the uterus of a pregnant sheep. The intelligibility scores were approximately 55% and 34% for the male and female talkers, respectively. However, the results from the current study showed that the intelligibility scores averaged across the stimulus types and intensity levels, were

approximately 91% and 85% for the male and female voices recorded in the uterus, respectively. The lower intelligibility of speech achieved by Querleu et al. (1988b) has been explained by the location of the transducer within the uterus (Griffiths et al. 1994) and by the type of transducer. A modified microphone used in Querleu's study was positioned at the crown of the fetal head, potentially closer to vascular beds and better able to pick up maternal heart sounds. In both the present study and the study by Griffiths et al. (1994), a hydrophone positioned by the fetal neck was used. The absence of detectable heart sounds in the recordings from these two studies supports that the hydrophone placement results in less vascular noise. However, the recordings within the uterus in the current study showed much higher intelligibility scores than that in Griffiths' study (1994), although both sets of data were obtained using the same speech stimuli (VCV nonsense syllable and CVC words) spoken by male and female speakers. The discrepancy could result from the higher stimulus levels (105 and 95 dB SPL vs. 85, 75, and 65 dB SPL) and better perceptual testing condition (earphone vs. sound field) used in the current study.

Griffiths et al. (1994) also demonstrated that the male talker's voice was more intelligible than the female talker's voice for both VCV and CVC stimuli when recorded within the uterus, although the intelligibility scores for both talkers were not significantly different regardless of stimulus type when recorded in air. The results of the present data indicated that the intelligibility scores of the male voice were significantly higher than that of the female voice across all four recording locations (in air, in the uterus, from CM *ex utero*, and from CM *in utero*) for VCV nonsense syllables, but not for CVC words. The differences of intelligibility scores for VCV nonsense syllables between the male

talker and the female talker were 8.1% in air (98.9% for male and 90.8% for female), 12.1% in the uterus (92.9% for male and 80.8% for female), 25.1% from CM *ex utero* (75.3% for male and 49.4% for female), and 5.6% from CM *in utero* (39.6% for male and 34.0% for female). When listening to the female speaker's original tape, it is difficult for investigators to distinguish the consonant /v/ from /b/. Twenty-nine out of 33 judges responded VCV stimulus item /ava/ as /aba/ in the air condition for the female talker. The unclear pronunciation of the consonant /v/ accounted for the decreases in the intelligibility of the female talker in air and, therefore, for the other recording locations.

The differences in the intelligibility scores between the male and the female talkers ranged from 5.6% (CM *in utero*) to 25.1% (CM *ex utero*). These differences were quite small (except 25.1% for CM *ex utero*) relative to the 14-item perceptual test (14 VCV items). Thus, the differences between talker gender may not be clinically significant, although they are statistically significant. Thornton and Raffin (1978) studied the binomial characteristics of speech discrimination (intelligibility) scores and pointed out the relation between measurement error and sample size (number of test item). As sample size was reduced, variability in scores increased, and the farther the score from 100% or 0% the less confidence one can have in the specific value. The authors have provided confidence intervals and expected ranges of scores based on evaluations of 4120 subjects with CID Auditory Test W-22 (monosyllabic words). For example, a listener who makes a score of 92% may vary between 78 and 98% on a 50-item list and still be within expected variation (95% confidence interval) while the expected range of variation for a 25-item list is even greater at 72 to 100%. For the subject with a score of 48%, the range of variation for 50 items is from 30 to 66%, and for 25 items is 24 to 72%.

The CM is an AC receptor potential produced primarily by the outer hair cells of the organ of Corti during acoustic stimulation, and mimics the acoustic input in amplitude and frequency over a remarkably wide range (Gulick, Gescheider and Frisina, 1989). In response to complex stimuli, like speech or music, the CM continues to follow the stimulus waveform, although there is some phase distortion due to the differing travel times necessary for the distribution of the various frequencies to their appropriate places along the cochlear basilar membrane. Nevertheless, when the CM is suitably amplified and converted back into sound, speech and music are easily recognizable.

In the current study, the recordings from the CM *ex utero* condition represented the actual fetal responses to speech in air that simulated the auditory condition of after birth. The CM *in utero* recordings reflected the speech information preserved in the fetal peripheral auditory system after transmission of external speech from air through the maternal tissues and fluids into the fetal inner ear. However, since CM is not an ideal "microphone," the overall intelligibility from CM *ex utero* recordings was only 61.9% when averaged across gender, intensity level, and stimulus type. Several factors can be accounted for this low intelligibility score. First, a high-level background noise was created by the biological amplifier used during CM recordings. This would decrease the S/N ratio and increase the difficulty of the perceptual test. Second, during the *ex utero* CM recordings, fluids might have been retained in the middle ear cavity and perhaps external ear canal, although special attention was paid to the clearing of these fluids. The retained fluid would increase the mass of the middle ear, which could reduce the transmission of high-frequency sounds into the inner ear (Pickles, 1988; Gulick, Gescheider and Frisina, 1989). Thus, high-frequency components of speech would be

attenuated in the recordings from CM *ex utero*, if fluid remained in the middle ear cavity. Finally, the CM produced by any particular pure tone has its maximum sensitivity at a specific place along the cochlear basilar membrane (Gulick, Gescheider and Frisina, 1989). Honrubia and Ward (1968) determined the spatial distribution of the CM inside the scala media by recording simultaneously from each of four electrodes in the scala media, one in each turn of the guinea pig cochlea. They found that the places of maximum CM shifted toward the basal end of the cochlea as the frequency of the driving stimulus increased. The spread of the electrical potential was less for the higher frequencies than for the lower frequencies, just as anticipated from the basis of the traveling wave. Therefore, the CM measured with a single electrode placed on the round window membrane, which was used in the present study, only accurately measured the response of hair cells from the basal turn, and cannot record the entire cochlear response to the input signals. The low-frequency information of the speech signal would be reduced in the CM recordings by using a single round window electrode. Overall, in the present study, the recordings of speech from CMs underestimated the speech information actually preserved in the fetal inner ear.

In summary, when the mean intelligibility scores were averaged across two stimulus levels (105 and 95 dB SPL) and stimulus types (VCV and CVC stimuli), they were 97.7% and 93.1% for the male and female voices recorded in air. Within the uterus, scores were 91.2% and 84.7% for the male and the female voices, respectively. The decline in intelligibility was only 6.5% for the male speaker and 8.4% for the female speaker from in air recordings to in the uterus recordings. The reduction of intelligibility reflected the filter effect produced by the maternal abdomen, uterus and amniotic fluid.



In contrast, the mean intelligibility scores recorded from CM *in utero*, averaged across two levels and stimulus types for the male and female voices, were 45.0% and 41.6%, respectively. For CM *ex utero*, scores were 67.7% and 56.1% for the male and the female talkers. Thus the reduction in intelligibility recordings made from CM *ex utero* to CM *in utero* was 22.7% for the male speaker and 14.5% for the female speaker, which was greater than that from in air recordings to in the uterus recordings. These declines in intelligibility represented the loss of speech information after transmission from air through the tissues and fluids associated with pregnancy and transmission through the fetal skull into the inner ear.

The results reported in this section support the hypotheses that the intelligibility of monosyllabic words and nonsense syllables will be reduced when recorded in the uterus compared to air. The results also explain the hypotheses that the intelligibility of monosyllabic words and nonsense syllables will be reduced when recorded from the fetal inner ear *in utero* compared to uterus.

Finally, the hypotheses were not supported that the intelligibility of a male talker will be greater than the intelligibility of a female talker when recorded in the uterus and from the fetal inner ear *in utero*. Two explanations for the lack of support for these hypotheses are offered. First, high level presentations (105 and 95 dB SPL) of the nonsense syllables and monosyllabic words may have produced improved intelligibility for the female talker when recorded in the uterus and from the fetal inner ear. And earlier study with sheep that showed a gender effect used lower levels of stimulus presentations (Griffiths et al., 1994). Second, in the CM *in utero* recording condition, the gender effect on the intelligibility of speech stimuli was minimized because the cutoff frequency of the

low-pass filter (bone conduction route into the fetal inner ear) was further lowered when compared to the uterus recording condition (Gerhardt et al., 1992). The high-frequency component of speech that cued the gender effect on the intelligibility of speech was eliminated when transmitted into the fetal inner ear *in utero*.

### Consonant Feature Transmission

Consonant confusion matrices were constructed from the responses of all subjects to the VCV nonsense syllables for each recording condition. They are presented in Tables 4-7 to 4-14 for the male talker, at each of the four recording locations (air, uterus, CM *ex utero*, and CM *in utero*), and at each of the stimulus levels (105 and 95 dB SPL). Tables 4-15 to 4-22 are the consonant confusion matrices for the female talker under the different recording conditions. In general, some consonant confusion patterns can be derived from inspecting these matrices. First, all 14 consonants were identified with high accuracy from recordings made in air. The accuracy of identification was slightly less in the uterus for both male and female speakers. In the recordings from CM *ex utero* and from CM *in utero*, accuracy of identification decreased dramatically.

Some consonants were correctly identified consistently across the recording conditions, while others were not. For example, the nasal /n/ was perceived quite accurately in all recording conditions, while correct identification of the fricative /s/ and the affricate /tʃ/ were much lower in the recordings from the fetal inner ear. The correct identification of the /n/ sound was 100% in the recordings made in air, in the uterus, and from CM *ex utero* for both male and female talkers at 105 dB SPL. It dropped slightly to 73% and 83% in CM *in utero* at 105 dB SPL for the male and female talkers,











Table 4-12. Consonant confusion matrix for male talker, recorded from CM-ex *utero* at 95 dB SPL.

| Stimulus | Response |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|----------|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
|          | /b/      | /p/ | /d/ | /t/ | /g/ | /k/ | /f/ | /v/ | /s/ | /z/ | /m/ | /n/ | /ʃ/ | /ʒ/ |  |
| /b/      | 33       | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /p/      | 0        | 30  | 0   | 2   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /d/      | 0        | 0   | 18  | 0   | 15  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /t/      | 0        | 0   | 0   | 18  | 0   | 15  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /g/      | 0        | 0   | 0   | 0   | 33  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /k/      | 0        | 0   | 0   | 1   | 0   | 31  | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   |  |
| /f/      | 0        | 0   | 0   | 10  | 0   | 9   | 9   | 0   | 3   | 0   | 0   | 0   | 1   | 1   |  |
| /v/      | 0        | 0   | 0   | 0   | 0   | 0   | 0   | 30  | 0   | 3   | 0   | 0   | 0   | 0   |  |
| /s/      | 0        | 0   | 0   | 2   | 2   | 14  | 6   | 0   | 7   | 1   | 0   | 0   | 1   | 0   |  |
| /z/      | 0        | 0   | 0   | 0   | 3   | 0   | 0   | 0   | 0   | 30  | 0   | 0   | 0   | 0   |  |
| /m/      | 0        | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 33  | 0   | 0   | 0   |  |
| /n/      | 0        | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 3   | 29  | 0   | 0   |  |
| /ʃ/      | 0        | 1   | 0   | 2   | 0   | 8   | 0   | 0   | 2   | 0   | 0   | 0   | 19  | 1   |  |
| /ʒ/      | 0        | 1   | 0   | 3   | 0   | 14  | 0   | 1   | 4   | 0   | 0   | 0   | 6   | 4   |  |



Table 4-13. Consonant confusion matrix for male talker, recorded from CM-in utero at 105 dB SPL.

| Stimulus | Response |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----------|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|          | /b/      | /p/ | /d/ | /t/ | /g/ | /k/ | /f/ | /v/ | /s/ | /z/ | /m/ | /n/ | /ʃ/ | /h/ |
| /b/      | 22       | 11  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /p/      | 0        | 4   | 0   | 18  | 0   | 7   | 4   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /d/      | 0        | 0   | 1   | 0   | 32  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /t/      | 0        | 3   | 0   | 4   | 0   | 26  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /g/      | 0        | 0   | 0   | 0   | 31  | 1   | 0   | 0   | 0   | 0   | 1   | 0   | 0   | 0   |
| /k/      | 0        | 1   | 0   | 5   | 0   | 27  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /f/      | 0        | 3   | 0   | 1   | 0   | 2   | 21  | 0   | 6   | 0   | 0   | 0   | 0   | 0   |
| /v/      | 0        | 0   | 0   | 0   | 0   | 0   | 0   | 27  | 0   | 6   | 0   | 0   | 0   | 0   |
| /s/      | 1        | 3   | 0   | 3   | 1   | 3   | 15  | 0   | 6   | 0   | 1   | 0   | 0   | 0   |
| /z/      | 0        | 0   | 0   | 0   | 9   | 0   | 0   | 6   | 0   | 18  | 0   | 0   | 0   | 0   |
| /m/      | 0        | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 27  | 6   | 0   | 0   |
| /n/      | 0        | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 9   | 24  | 0   | 0   |
| /ʃ/      | 0        | 0   | 0   | 1   | 0   | 1   | 6   | 0   | 21  | 0   | 0   | 0   | 4   | 0   |
| /h/      | 0        | 0   | 0   | 15  | 0   | 4   | 0   | 0   | 10  | 0   | 0   | 0   | 4   | 0   |

Table 4-14. Consonant confusion matrix for male talker, recorded from CM-in utero at 95 dB SPL.

| Stimulus | Response |     |     |     |     |     |      |     |     |     |     |     |     |     |     |     |
|----------|----------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|          | /b/      | /p/ | /d/ | /t/ | /g/ | /k/ | /tʃ/ | /ʃ/ | /s/ | /z/ | /m/ | /n/ | /ʃ/ | /s/ | /h/ | /ʃ/ |
| /b/      | 28       | 0   | 2   | 2   | 0   | 0   | 0    | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   |
| /p/      | 3        | 3   | 0   | 1   | 0   | 6   | 7    | 0   | 4   | 1   | 0   | 0   | 8   | 0   | 0   | 0   |
| /d/      | 0        | 0   | 1   | 0   | 25  | 0   | 0    | 6   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| /t/      | 1        | 6   | 0   | 2   | 1   | 15  | 6    | 0   | 0   | 0   | 0   | 0   | 2   | 0   | 0   | 0   |
| /g/      | 2        | 2   | 3   | 3   | 18  | 2   | 0    | 1   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   |
| /k/      | 0        | 1   | 0   | 13  | 0   | 11  | 4    | 0   | 0   | 0   | 0   | 0   | 1   | 3   | 0   | 0   |
| /tʃ/     | 0        | 0   | 0   | 4   | 0   | 4   | 6    | 0   | 8   | 0   | 0   | 0   | 4   | 7   | 0   | 0   |
| /ʃ/      | 3        | 0   | 2   | 0   | 9   | 0   | 0    | 16  | 0   | 3   | 0   | 0   | 0   | 0   | 0   | 0   |
| /s/      | 1        | 3   | 0   | 5   | 0   | 4   | 11   | 0   | 4   | 0   | 0   | 0   | 3   | 2   | 0   | 0   |
| /z/      | 0        | 0   | 4   | 0   | 13  | 0   | 0    | 0   | 0   | 15  | 0   | 0   | 1   | 0   | 0   | 0   |
| /m/      | 0        | 0   | 0   | 0   | 0   | 0   | 0    | 0   | 0   | 0   | 32  | 1   | 0   | 0   | 0   | 0   |
| /n/      | 0        | 0   | 0   | 0   | 0   | 0   | 0    | 0   | 0   | 0   | 20  | 13  | 0   | 0   | 0   | 0   |
| /ʃ/      | 2        | 11  | 0   | 0   | 4   | 5   | 10   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /tʃ/     | 0        | 0   | 0   | 3   | 0   | 9   | 8    | 1   | 6   | 1   | 0   | 0   | 4   | 1   | 0   | 0   |









Table 4-19. Consonant confusion matrix for female talker, recorded from CM-ex *utero* at 105 dB SPL.

| Stimulus | Response |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|----------|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
|          | /b/      | /p/ | /d/ | /t/ | /g/ | /k/ | /f/ | /v/ | /s/ | /z/ | /m/ | /n/ | /ʃ/ | /ʒ/ |  |
| /b/      | 30       | 0   | 2   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /p/      | 0        | 19  | 0   | 6   | 0   | 8   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /d/      | 1        | 1   | 22  | 0   | 9   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /t/      | 0        | 4   | 0   | 22  | 0   | 6   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   |  |
| /g/      | 4        | 0   | 7   | 0   | 22  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /k/      | 0        | 0   | 0   | 3   | 0   | 30  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /f/      | 0        | 11  | 0   | 4   | 0   | 1   | 11  | 0   | 2   | 0   | 0   | 0   | 3   | 1   |  |
| /v/      | 20       | 0   | 5   | 0   | 8   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |  |
| /s/      | 0        | 1   | 2   | 6   | 2   | 18  | 1   | 0   | 2   | 1   | 0   | 0   | 0   | 0   |  |
| /z/      | 0        | 0   | 19  | 0   | 12  | 0   | 0   | 0   | 0   | 2   | 0   | 0   | 0   | 0   |  |
| /m/      | 0        | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 33  | 0   | 0   | 0   |  |
| /n/      | 0        | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 33  | 0   | 0   |  |
| /ʃ/      | 0        | 3   | 2   | 11  | 2   | 3   | 2   | 0   | 4   | 4   | 0   | 0   | 1   | 1   |  |
| /ʒ/      | 0        | 1   | 3   | 15  | 0   | 7   | 1   | 0   | 1   | 0   | 0   | 0   | 0   | 5   |  |





Table 4-21. Consonant confusion matrix for female talker, recorded from CM-in utero at 105 dB SPL.

| Stimulus | Response |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----------|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|          | /b/      | /p/ | /d/ | /t/ | /g/ | /k/ | /f/ | /v/ | /s/ | /z/ | /m/ | /n/ | /ŋ/ | /ʃ/ | /ʒ/ |
| /b/      | 16       | 0   | 12  | 0   | 5   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /p/      | 1        | 17  | 0   | 3   | 1   | 8   | 3   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /d/      | 0        | 0   | 20  | 0   | 13  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /t/      | 0        | 5   | 0   | 19  | 0   | 9   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /g/      | 1        | 0   | 13  | 0   | 19  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /k/      | 0        | 7   | 0   | 11  | 0   | 10  | 4   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   |
| /f/      | 0        | 9   | 0   | 1   | 1   | 5   | 11  | 0   | 2   | 0   | 0   | 0   | 3   | 1   | 0   |
| /v/      | 11       | 0   | 0   | 1   | 21  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /s/      | 0        | 6   | 0   | 0   | 1   | 6   | 11  | 0   | 2   | 2   | 0   | 0   | 5   | 0   | 0   |
| /z/      | 0        | 0   | 4   | 0   | 8   | 0   | 0   | 0   | 0   | 21  | 0   | 0   | 0   | 0   | 0   |
| /m/      | 0        | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 3   | 30  | 0   | 0   | 0   |
| /n/      | 0        | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 6   | 27  | 0   | 0   | 0   |
| /ŋ/      | 1        | 3   | 0   | 6   | 1   | 1   | 15  | 0   | 0   | 3   | 0   | 0   | 2   | 1   | 0   |
| /ʃ/      | 0        | 8   | 0   | 3   | 0   | 0   | 2   | 0   | 0   | 0   | 0   | 0   | 5   | 15  | 0   |

Table 4-22. Consonant confusion matrix for female talker, recorded from CM-in utero at 95 dB SPL.

| Stimulus | Response |     |     |     |     |     |      |     |     |     |     |     |     |     |
|----------|----------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|
|          | /b/      | /p/ | /d/ | /t/ | /g/ | /k/ | /tʃ/ | /v/ | /s/ | /z/ | /m/ | /n/ | /ʃ/ | /ʒ/ |
| /b/      | 18       | 0   | 0   | 0   | 7   | 1   | 0    | 6   | 0   | 0   | 1   | 0   | 0   | 0   |
| /p/      | 0        | 15  | 0   | 0   | 0   | 11  | 6    | 0   | 1   | 0   | 0   | 0   | 0   | 0   |
| /d/      | 7        | 0   | 5   | 0   | 16  | 0   | 1    | 0   | 0   | 4   | 0   | 0   | 0   | 0   |
| /t/      | 0        | 5   | 0   | 19  | 0   | 3   | 2    | 0   | 0   | 0   | 0   | 0   | 2   | 2   |
| /g/      | 5        | 0   | 6   | 0   | 19  | 0   | 1    | 2   | 0   | 0   | 0   | 0   | 0   | 0   |
| /k/      | 0        | 5   | 0   | 4   | 7   | 15  | 1    | 0   | 0   | 0   | 0   | 1   | 0   | 0   |
| /tʃ/     | 0        | 9   | 0   | 2   | 0   | 7   | 11   | 0   | 1   | 0   | 0   | 0   | 1   | 2   |
| /v/      | 10       | 0   | 4   | 0   | 19  | 0   | 0    | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| /s/      | 6        | 5   | 2   | 0   | 7   | 6   | 5    | 0   | 1   | 1   | 0   | 0   | 0   | 0   |
| /z/      | 2        | 0   | 2   | 0   | 16  | 1   | 0    | 0   | 1   | 8   | 0   | 0   | 3   | 0   |
| /m/      | 0        | 0   | 0   | 0   | 1   | 0   | 0    | 0   | 0   | 0   | 9   | 23  | 0   | 0   |
| /n/      | 0        | 0   | 0   | 0   | 0   | 0   | 0    | 0   | 0   | 1   | 28  | 4   | 0   | 0   |
| /ʃ/      | 2        | 1   | 3   | 3   | 6   | 6   | 0    | 0   | 3   | 2   | 3   | 0   | 3   | 1   |
| /ʒ/      | 1        | 8   | 0   | 3   | 2   | 6   | 2    | 0   | 1   | 0   | 0   | 0   | 5   | 5   |

respectively. In contrast, for the male voice recorded at 105 dB SPL, correct identifications of /S/ and /tS/ were 85% and 36% from CM *ex utero*, and 12% and 0% from CM *in utero*, respectively, although both consonants were perfectly identified (100%) in air and in the uterus conditions. For the female voice recorded at 105 dB SPL, correct identification of /S/ and /tS/ were 3% and 15% from CM *ex utero*, and 6% and 45% from CM *in utero*, respectively, while /S/ was 88% identified, and /tS/ was 100% identified both in air and in the uterus conditions. Further analyses of the consonant feature transmission under different recording conditions were made using a special computer program.

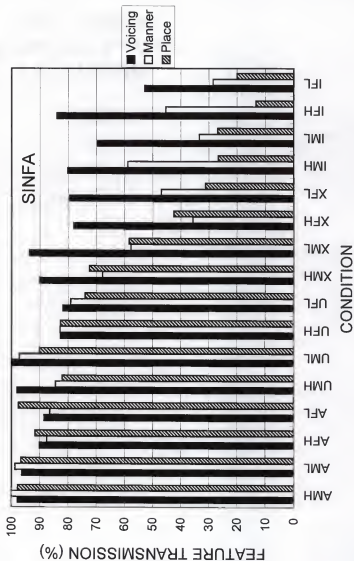
Because the features of voicing, manner, and place are strongly interdependent, the sequential information analysis (SINFA), which sequentially identifies features with a high proportion of transmitted information, was applied to partial out the effects of the features on each other (Wang and Bilger, 1973; Wang, 1976). SINFA focuses on the transmitted information associated with a given stimulus-response confusion matrix and identifies the contributions of various phonological features to the transmitted information. The results of SINFA are given in Table 4-23, which contains the percentage of contingent voicing, manner, and place information received (bits received / bits sent) for each talker and recording location for 105-dB and 95-dB stimuli. The SINFA results are graphically displayed in Figure 4-5.

A number of points can be made from inspection of Figure 4-5. First, all three features, voicing, manner and place, appeared to be well transmitted in recordings made in air regardless of talker gender and stimulus levels. Second, voicing information received from in the uterus recordings was slightly reduced, about 7%, for the female talker but not at

Table 4-23. Conditional percentage of voicing, manner, and place information received (of bits sent) for each talker, recording location, and stimulus level condition for the nonsense syllables (VCV).

| Location    | Condition |      |        |      |           |      |        |      |             |      |        |      |             |      |        |      |
|-------------|-----------|------|--------|------|-----------|------|--------|------|-------------|------|--------|------|-------------|------|--------|------|
|             | In Air    |      |        |      | In Uterus |      |        |      | CM-ex utero |      |        |      | CM-in utero |      |        |      |
|             | Male      |      | Female |      | Male      |      | Female |      | Male        |      | Female |      | Male        |      | Female |      |
| Talker      |           |      |        |      |           |      |        |      |             |      |        |      |             |      |        |      |
| Level (dB)  | 105       | 95   | 105    | 95   | 105       | 95   | 105    | 95   | 105         | 95   | 105    | 95   | 105         | 95   | 105    | 95   |
| Information |           |      |        |      |           |      |        |      |             |      |        |      |             |      |        |      |
| Voicing     | 97.9      | 96.4 | 90.2   | 88.5 | 98.0      | 100  | 82.7   | 81.8 | 90.1        | 93.7 | 78.0   | 79.5 | 80.2        | 69.8 | 84.0   | 52.8 |
| Manner      | 100       | 98.7 | 87.5   | 86.5 | 84.4      | 97.2 | 82.8   | 79.1 | 67.7        | 57.7 | 35.7   | 46.9 | 58.8        | 33.4 | 45.3   | 28.5 |
| Place       | 97.8      | 96.7 | 91.7   | 97.5 | 82.0      | 90.1 | 82.6   | 73.9 | 72.4        | 58.3 | 42.4   | 31.2 | 26.7        | 26.8 | 13.3   | 20.0 |

Figure 4-5. Conditional percentage of voicing, manner and place information received for a male (M) and a female (F) talker; in air (A), in the uterus (U), from the fetal CM *ex utero* (X), and from the fetal CM *in utero* (I); at 105 dB (H) and 95 dB (L) SPL.



all for the male talker across the stimulus levels. However, manner and place information were reduced about 6-8% and 10-15%, respectively, for both male and female talkers when recorded in the uterus and averaged across stimulus levels. Third, information about all three features decreased from hydrophone recordings within the uterus to recordings from fetal CM *ex utero*, and to that from CM *in utero*. However, voicing information appeared to be better preserved than manner and place information for both male and female talkers. Voicing information received from CM *ex utero* and CM *in utero* was ranged from 94% to 70% for the male talker, and from 84% to 53% for the female talker. Information about manner and place was reduced markedly in CM *ex utero* and CM *in utero* recordings, especially for the female talker. For CM *in utero* recordings, manner information was reduced less than place information. In all cases of feature information received from CM recordings, there was a greater loss of each of the three features information for the female speaker than for the male speaker, except for voicing information received from CM *in utero* at 105 dB SPL.

In the previous study conducted by Griffiths et al. (1994), a panel of 102 untrained individuals judged the intelligibility of speech recorded *in utero* from a pregnant sheep. The same VCV and CVC stimuli were used as the present study. An analysis (SINFA) of the feature information from recordings inside and outside the uterus showed that voicing information is better transmitted *in utero* than place or manner information. The current study confirmed the findings regarding voicing information inside the uterus. Furthermore, the results of SINFA from the present study indicated that voicing information was accurately perceived in the fetal inner ear (CM recordings) *ex utero* and *in utero*, and the male voicing information was better preserved than that of the

female. Manner and place information were not received as well as voicing information by the fetal inner ear; there were remarkable reductions in CM recordings, especially for the female voice.

Miller and Nicely (1955) reported that low-pass filtering of speech signals resulted in a greater loss of manner and place information than of voicing information. They concluded that the higher frequency information in the speech signal is critical for accurate identification of manner and place of articulation. Wang et al. (1978) had the same conclusion on consonant feature recognition of low-pass filtering speech by using SINFA.

The findings of both Griffiths et al. (1994) and the current study are consistent with those of Miller and Nicely (1955) and Wang et al. (1978) in that transmission into the uterus can be modeled as a low-pass filter. The poorer *in utero* reception of place and manner information is associated with the greater high-frequency attenuation. Moreover, the spectral contents of external speech signals are further modified by the route of bone conduction through the fetal skull to the inner ear (Gerhardt et al., 1996). For low frequencies, 125 and 250 Hz, an airborne signal would be reduced by 10-20 dB to reach the fetal inner ear *in utero*. For 500 through 2000 Hz, the signal would be reduced by 35-45 dB (Gerhardt et al., 1992). Thus, the high-frequency components of speech would be attenuated once again when transmitted through the skull into the fetal inner ear *in utero*. Manner and place information were lost to a great degree in the recordings from CM *in utero*, since high-frequency information was attenuated most after transmission from air through the maternal abdomen, uterus, and fetal head to the fetal inner ear.



The results derived from SINFA support the hypotheses that transmission into the uterus and fetal inner ear will be greater for voicing information than for manner and place information. The results further support the hypotheses that the transmission of voicing, manner, and place information will be better for males than for females when recorded in the uterus and from the inner ear of the fetus *in utero*.

Voicing information from the male talker, which is carried by low-frequency energy, was largely preserved inside the uterus and also in the fetal inner ear *in utero*. The judges evaluated the male talker's voice equally well regardless of recording location. Speech of the female talker carried less well into the uterus and into the fetal inner ear *in utero*. The fundamental frequency of the female talker was about an octave higher than that of the male talker. Thus, it is predictable that voicing information from the male would carry better into the uterus, and into the fetal inner ear *in utero* than that from the female.

#### Acoustic Analyses of Vowel Transmission

Figure 4-6 A-H includes sample spectrographs displaying one stimulus item recorded in eight conditions: the male and female talkers recorded in air, in the uterus, from CM *ex utero* and from CM *in utero* at 105 dB SPL. The phrase spoken in each of the eight spectrographs is "Mark the word lash." The amplitudes of each recording were adjusted to the same relative voltage level on the spectrographic analysis. The contrast between voiced and voiceless portions of the phrase is apparent to some degree in all eight spectrographs. The high-frequency noise associated with the release of the fricative,

Figure 4-6. Spectrographic recordings of "Mark the word lash" spoken by: A. the male talker recorded in air; B. the female talker recorded in air; C. the male talker recorded in the uterus; D. the female talker recorded in the uterus; E. the male talker recorded from the fetal CM *ex utero*; F. the female talker recorded from the fetal CM *ex utero*; G. the male talker recorded from the fetal CM *in utero*; H. the female talker recorded from the fetal CM *in utero*.

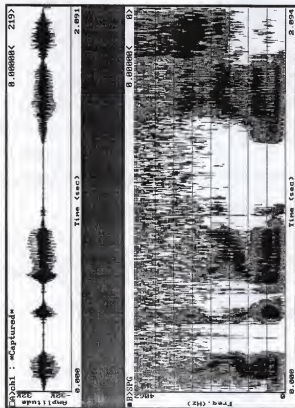


Figure 4-6 A.

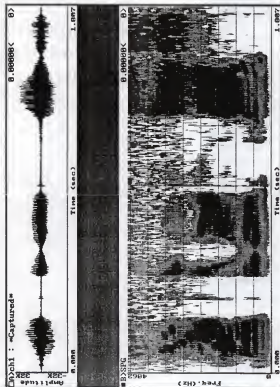


Figure 4-6 B. Continued.

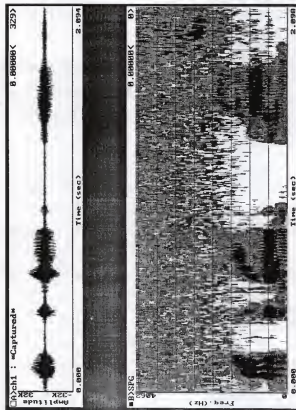


Figure 4-6 C. Continued.

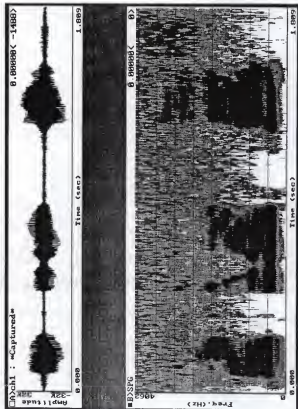


Figure 4-6 D. Continued.

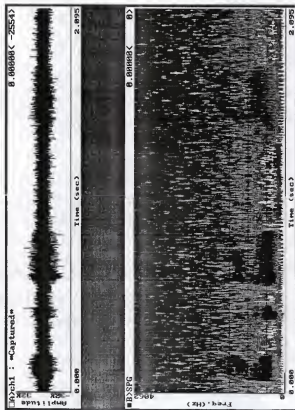


Figure 4-6 E. Continued

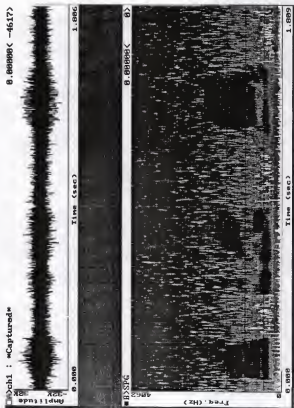


Figure 4-6 F. Continued.



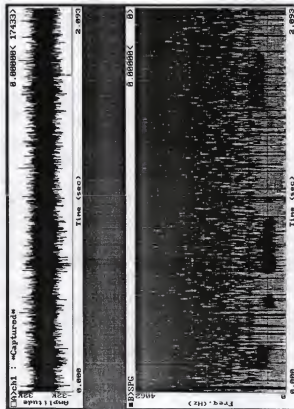


Figure 4-6 G. Continued.

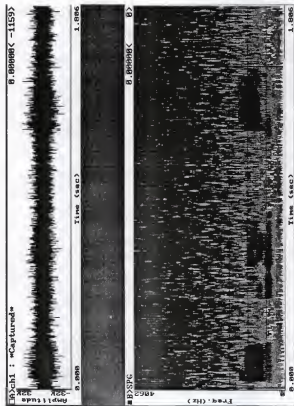


Figure 4-6 H. Continued.

/S/, is undetectable in the CM *in utero* spectrograms for both talkers, consistent with the low-pass filtering of the maternal tissues, fluids, and fetal skull.

Acoustic measurements for the CVC words containing one of the five vowels (/i/, /I/, /e/, /æ/, /ʌ/) were performed. For each of the vowels /i/, /I/, /æ/, and /ʌ/, five CVC words were selected for spectral analyses. For the vowel /e/, four CVC words were analyzed. The means of the fundamental frequency ( $F_0$ ) and the first three formant frequencies ( $F_1$ ,  $F_2$ , and  $F_3$ ) of the male and female speakers averaged across each of the five vowels are presented in Table 4-24. For the purpose of comparison, Table 4-24 also includes the values obtained from two large studies of vowel formant frequencies, a classic paper by Peterson and Barney (1952) and a recent replication by Hillenbrand et al. (1995). There were clear similarities in the present data to the data from Peterson and Barney (1952), and from Hillenbrand et al. (1995).

To evaluate the characteristics of vowel transmission under different recording conditions, the relative intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were calculated by subtracting the background noise level from the peak amplitudes of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  for both the male and female speakers across the five vowels. Table 4-25 contains the means and standard deviations of relative intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  for vowel /i/ produced by the male and female talkers under different recording conditions. These data are also displayed in Figure 4-7. Vowel /i/ has a low  $F_1$  frequency (345 Hz for the male speaker and 353 Hz for female) and a high  $F_2$  frequency (2490 Hz for male and 2841 Hz for female), as well as a high  $F_3$  frequency (3590 Hz for male and 3337 Hz for female). From an inspection of Figure 4-7, a general transmission pattern for vowel /i/ can be drawn. In the air recording condition,  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were well identified for both male and female talkers. For the

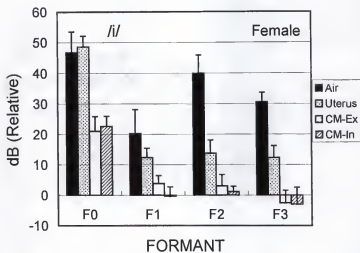
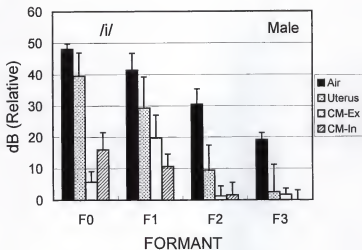
Table 4-24. Average fundamental frequencies ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ ,  $F_3$ ) for five vowels produced by each talker and recorded in air. The second row includes the value from Peterson and Barney (1952). The third row is from Hillenbrand et al. (1995).

|       |        | /i/  | /I/  | /e/  | /æ/  | /A/  |
|-------|--------|------|------|------|------|------|
| $F_0$ | Male   | 136  | 123  | 128  | 113  | 118  |
|       | (1952) | 136  | 135  | 130  | 127  | 130  |
|       | (1995) | 138  | 135  | 127  | 123  | 133  |
|       | Female | 242  | 224  | 220  | 217  | 221  |
|       | (1952) | 275  | 232  | 223  | 210  | 221  |
|       | (1995) | 270  | 224  | 214  | 215  | 218  |
| $F_1$ | Male   | 345  | 397  | 584  | 666  | 622  |
|       | (1952) | 270  | 390  | 530  | 660  | 640  |
|       | (1995) | 342  | 427  | 580  | 588  | 623  |
|       | Female | 353  | 442  | 650  | 1069 | 670  |
|       | (1952) | 310  | 430  | 610  | 860  | 760  |
|       | (1995) | 437  | 483  | 731  | 669  | 753  |
| $F_2$ | Male   | 2490 | 2076 | 1831 | 1676 | 1242 |
|       | (1952) | 2290 | 1990 | 1840 | 1720 | 1190 |
|       | (1995) | 2322 | 2034 | 1799 | 1952 | 1200 |
|       | Female | 2841 | 2376 | 1975 | 1926 | 1278 |
|       | (1952) | 2790 | 2480 | 2330 | 2050 | 1400 |
|       | (1995) | 2761 | 2365 | 2058 | 2349 | 1426 |
| $F_3$ | Male   | 3597 | 2690 | 2766 | 2567 | 2696 |
|       | (1952) | 3010 | 2550 | 2480 | 2410 | 2390 |
|       | (1995) | 3000 | 2684 | 2605 | 2601 | 2550 |
|       | Female | 3337 | 2889 | 2845 | 2904 | 2789 |
|       | (1952) | 3310 | 3070 | 2990 | 2850 | 2780 |
|       | (1995) | 3372 | 3053 | 2979 | 2972 | 2933 |

Table 4-25. Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ ,  $F_3$ ) for vowel /i/ produced by each talker at different recording locations in the 105 dB condition.

| Male talker, /i/   |      | Condition |           |             |             |
|--------------------|------|-----------|-----------|-------------|-------------|
|                    |      | In Air    | In Uterus | CM-ex utero | CM-in utero |
| $F_0$              | Mean | 48.12     | 39.51     | 5.70        | 16.00       |
|                    | S.D. | 1.69      | 7.40      | 3.40        | 5.51        |
| $F_1$              | Mean | 41.38     | 29.30     | 19.78       | 10.68       |
|                    | S.D. | 5.32      | 9.95      | 7.24        | 3.97        |
| $F_2$              | Mean | 30.50     | 9.50      | 1.20        | 1.54        |
|                    | S.D. | 4.92      | 7.93      | 3.20        | 4.04        |
| $F_3$              | Mean | 19.20     | 2.52      | 1.70        | 0.02        |
|                    | S.D. | 2.20      | 8.75      | 1.94        | 3.13        |
| Female talker, /i/ |      |           |           |             |             |
| $F_0$              | Mean | 46.82     | 48.64     | 20.98       | 22.60       |
|                    | S.D. | 6.72      | 3.51      | 4.88        | 3.32        |
| $F_1$              | Mean | 20.32     | 12.36     | 3.76        | -0.40       |
|                    | S.D. | 7.80      | 3.09      | 2.69        | 3.12        |
| $F_2$              | Mean | 39.92     | 13.72     | 3.00        | 1.08        |
|                    | S.D. | 5.96      | 4.38      | 3.70        | 1.76        |
| $F_3$              | Mean | 30.54     | 12.28     | -2.52       | -2.94       |
|                    | S.D. | 3.13      | 3.91      | 4.02        | 5.43        |

Figure 4-7. Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ , and  $F_3$ ) for vowel /i/ produced by both talkers recorded at different locations at 105 dB SPL. Bars equal one standard deviation. Male talker results - upper panel; female talker results - lower panel.



female, the intensity levels of  $F_2$  and  $F_3$  were 10 dB higher than that of male. However,  $F_1$  of the male talker was 20 dB greater than that of the female.

Now, considering the levels in the uterus,  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were well represented for both talkers, with the exception that  $F_3$  for the male talker was only about 3 dB above noise floor. The drop in levels as a function of formant frequency is predicted based upon transmission loss at higher frequencies.

From CM *ex utero* and *in utero* recordings,  $F_0$  and  $F_1$  for male and  $F_0$  only for female were preserved,  $F_2$  and  $F_3$  ( $F_1$  also for female) merged in the background noise. It was also noted that the intensity level of  $F_0$  from CM *in utero* was greater than that from CM *ex utero* for both talkers, especially for the male talker. The explanation is that low-frequency signals, less than 250 Hz ( $F_0$  were 136 Hz for the male talker and 242 Hz for female), would be enhanced when transmitted into the uterus (Vince et al., 1982; Gerhardt, Abrams and Oliver, 1990). Similar enhancement of  $F_0$  in the recordings from CM *in utero* was noted in other vowel measurements. Additionally, for the male talker the intensity levels of  $F_0$  were lower than female in the CM recordings, although  $F_0$  was equal intense in air. The CM measured by using a single round window electrode can not accurately record cochlear responses to the low-frequency input signals. Because the male  $F_0$  was one octave lower than female, it is understandable that the male  $F_0$  would be less detected than female by round window electrode. Less intense male  $F_0$  in CM recordings was noted in all five vowels. Thus, /i/ could be easily recognized in the uterus recordings for both talkers; however, its identification might not be made from CM *ex utero* and *in utero* recordings for the male talker because  $F_2$  was not perceived, and definitely not for the female talker because both  $F_1$  and  $F_2$  were not perceived.



Two separate, two-factor repeated measures ANOVA were applied to the data derived from vowel /i/ according to the talker gender. For the male talker, the results indicated significant interaction between formant ( $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$ ) and recording location ( $F_{9,36} = 7.334$ ,  $p < 0.0001$ ), main effects for formant ( $F_{3,12} = 43.543$ ,  $p < 0.0001$ ) and location ( $F_{3,12} = 203.061$ ,  $p < 0.0001$ ). The post hoc multiple comparison test (Newman-Keuls) indicated that the intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$  and  $F_3$  measured in air condition were significantly greater ( $p < 0.01$ ) than those recorded in the uterus, from CM *ex utero*, and from CM *in utero*, except that of  $F_0$  measured in the uterus ( $p > 0.05$ ). In the uterus condition, the intensity levels of  $F_0$  and  $F_1$  were greater ( $p < 0.01$ ; for  $F_1$  from CM *ex utero*  $p < 0.05$ ) than that from CM *ex utero* and from CM *in utero*, but not  $F_2$  and  $F_3$  ( $p > 0.05$ ). In the CM conditions, only  $F_0$  from CM *ex utero* was significantly different ( $p < 0.05$ ) from CM *in utero*; there were no difference ( $p > 0.05$ ) for the first three formant frequencies.

For the female talker, the results indicated significant interaction between formant ( $F_0$ ,  $F_1$ ,  $F_2$  and  $F_3$ ) and recording location ( $F_{9,36} = 13.49$ ,  $p < 0.0001$ ), main effects for formant ( $F_{3,12} = 87.767$ ,  $p < 0.0001$ ) and location ( $F_{3,12} = 117.211$ ,  $p < 0.0001$ ). The post hoc multiple comparison test (Newman-Keuls) indicated that the intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  measured in air condition were significantly greater ( $p < 0.01$ ) than those recorded in the uterus, from CM *ex utero*, and from CM *in utero*, except that of  $F_0$  measured in the uterus ( $p > 0.05$ ). In the uterus recording condition, the intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were greater ( $p < 0.01$ ) than that from CM *ex utero* and from CM *in utero*. Comparing conditions of CM *ex utero* and CM *in utero*, there were no difference ( $p > 0.05$ ) for the intensity levels of the fundamental frequency and the first three formant frequencies.

The explanation for the lack of significant differences is because most levels for  $F_2$  and  $F_3$  were indistinguishable from the noise floor.

Table 4-26 contains the means and standard deviations of relative intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  recorded in different locations for the vowel /l/ produced by the male and female talkers. The data are also displayed in Figure 4-8. Vowel /l/, similar to vowel /i/, has a low  $F_1$  frequency (397 Hz for the male speaker and 442 Hz for female) and a high  $F_2$  frequency (2076 Hz for male and 2376 Hz for female), as well as a high  $F_3$  frequency (2690 Hz for male and 2889 Hz for female), fairly close to  $F_2$  frequency. The results from spectral analyses and statistical analyses (ANOVA) were quite similar to that for vowel /i/. In the uterus,  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were also well received for both talkers. From CM *ex utero* and *in utero* recordings,  $F_0$  and  $F_1$  for both talkers were preserved, but  $F_2$  and  $F_3$  were less than 5 dB above the background noise. Therefore, /l/, like /i/, could be easily identified in the uterus recordings for both talkers, however, its identification might not be made from CM *ex utero* and *in utero* recordings for both talkers, because  $F_2$  was not well perceived.

Table 4-27 contains the means and standard deviations of relative intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  recorded in different locations for the vowel /e/ produced by the male and female talkers. The data are also displayed in Figure 4-9. In contrast to vowels /i/ and /l/, vowel /e/ has a high  $F_1$  frequency (584 Hz for the male speaker and 650 Hz for female) and a relative low  $F_2$  frequency (1831 Hz for male and 1975 Hz for female), as well as a high  $F_3$  frequency (2766 Hz for male and 2845 Hz for female). From an inspection of Figure 4-9, a general characteristic of transmission for the vowel /e/ can be derived. In the air recording condition,  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were well identified for both male and female talkers, and for the female the intensity levels of  $F_2$  and  $F_3$  were about 10 dB and 5 dB higher than that of

Table 4-26. Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ ,  $F_3$ ) for vowel /I/ produced by each talker at different recording locations in the 105 dB condition.

| Male talker, /I/   |      | Condition |           |             |             |
|--------------------|------|-----------|-----------|-------------|-------------|
|                    |      | In Air    | In Uterus | CM-ex utero | CM-in utero |
| $F_0$              | Mean | 51.64     | 45.50     | 6.18        | 11.04       |
|                    | S.D. | 5.29      | 5.62      | 3.04        | 2.51        |
| $F_1$              | Mean | 53.28     | 48.96     | 33.12       | 22.82       |
|                    | S.D. | 5.54      | 8.04      | 4.32        | 8.42        |
| $F_2$              | Mean | 36.90     | 19.40     | 0.48        | 1.66        |
|                    | S.D. | 6.93      | 9.77      | 2.71        | 3.51        |
| $F_3$              | Mean | 39.34     | 10.80     | 3.68        | 3.96        |
|                    | S.D. | 3.90      | 10.60     | 9.06        | 2.15        |
| Female talker, /I/ |      |           |           |             |             |
| $F_0$              | Mean | 44.64     | 39.88     | 18.10       | 20.30       |
|                    | S.D. | 4.23      | 4.21      | 3.09        | 5.50        |
| $F_1$              | Mean | 50.48     | 51.44     | 29.18       | 26.76       |
|                    | S.D. | 8.49      | 2.85      | 9.17        | 5.48        |
| $F_2$              | Mean | 37.76     | 21.62     | 3.08        | 1.22        |
|                    | S.D. | 9.87      | 4.98      | 4.86        | 4.41        |
| $F_3$              | Mean | 34.40     | 11.74     | 2.58        | -0.82       |
|                    | S.D. | 9.29      | 8.35      | 1.96        | 4.59        |

Figure 4-8. Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ , and  $F_3$ ) for vowel /l/ produced by both talkers recorded at different locations at 105 dB SPL. Bars equal one standard deviation. Male talker results - upper panel; female talker results - lower panel.

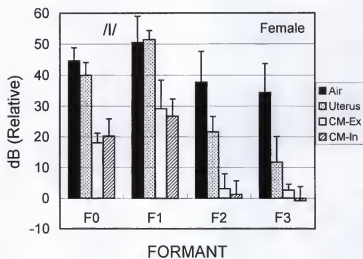
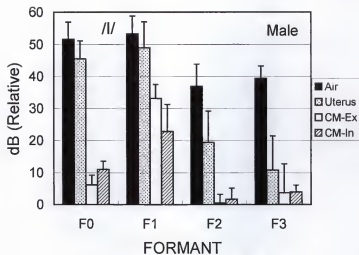
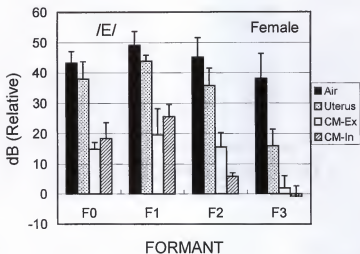
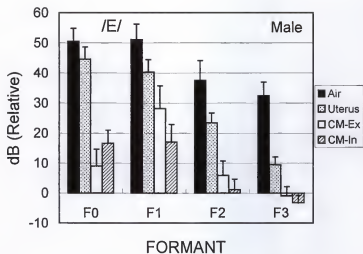


Table 4-27. Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ ,  $F_3$ ) for vowel / $\epsilon$ / produced by each talker at different recording locations in the 105 dB condition.

| Male talker, / $\epsilon$ /   |      | Condition |           |             |             |
|-------------------------------|------|-----------|-----------|-------------|-------------|
|                               |      | In Air    | In Uterus | CM-ex utero | CM-in utero |
| $F_0$                         | Mean | 50.53     | 44.60     | 8.98        | 16.65       |
|                               | S.D. | 4.24      | 4.04      | 5.62        | 4.33        |
| $F_1$                         | Mean | 51.10     | 40.28     | 28.15       | 17.08       |
|                               | S.D. | 5.10      | 4.18      | 7.60        | 5.82        |
| $F_2$                         | Mean | 37.58     | 23.40     | 5.83        | 1.15        |
|                               | S.D. | 6.56      | 3.24      | 4.89        | 3.47        |
| $F_3$                         | Mean | 32.48     | 9.38      | -0.95       | -3.23       |
|                               | S.D. | 4.46      | 2.70      | 3.11        | 2.41        |
| Female talker, / $\epsilon$ / |      |           |           |             |             |
| $F_0$                         | Mean | 43.25     | 37.98     | 14.85       | 18.40       |
|                               | S.D. | 3.83      | 5.73      | 2.28        | 5.16        |
| $F_1$                         | Mean | 49.15     | 43.83     | 19.58       | 25.55       |
|                               | S.D. | 4.52      | 1.96      | 8.59        | 4.05        |
| $F_2$                         | Mean | 45.18     | 35.80     | 15.55       | 5.83        |
|                               | S.D. | 6.41      | 5.74      | 4.71        | 1.15        |
| $F_3$                         | Mean | 38.20     | 15.90     | 1.90        | -0.85       |
|                               | S.D. | 8.14      | 5.50      | 4.12        | 3.49        |

Figure 4-9. Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ , and  $F_3$ ) for vowel /E/ (= /e/) produced by both talkers recorded at different locations at 105 dB SPL. Bars equal one standard deviation. Male talker results - upper panel; female talker results - lower panel.





the male, respectively. The female talker's higher intensity levels of  $F_2$  and  $F_3$  than the male in air resulted in higher levels of  $F_2$  and  $F_3$  measured in the uterus and fetal inner ear. In the uterus,  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were also well received for both talkers. From CM *ex utero* recordings,  $F_0$ ,  $F_1$ , and  $F_2$  for both talkers were transmitted into the fetal inner ear, but  $F_3$  merged in the background noise. From CM *in utero* recordings,  $F_0$ ,  $F_1$ , and  $F_2$  were preserved for the female talker, but only  $F_0$  and  $F_1$  were received for the male talker, since  $F_2$  was close to the level of background noise. Thus, /e/ could be easily identified in the uterus recordings for both talkers, and could be recognized from CM *ex utero* because  $F_2$  was well perceived for both talkers. However, its identification might be made from CM *in utero* recordings for the female talker, but might not be for the male because  $F_2$  was not well perceived in the fetal inner ear *in utero*.

Two separate, two-factor repeated measures ANOVA were applied to the data for the vowel /e/ for both the male and female talker. For the male talker, the results indicated significant interaction between formant ( $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$ ) and recording location ( $F_{9,27} = 6.612$ ,  $p < 0.0001$ ), main effects for formant ( $F_{3,9} = 68.103$ ,  $p < 0.0001$ ) and location ( $F_{3,9} = 163.051$ ,  $p < 0.0001$ ). The post hoc multiple comparison test (Newman-Keuls) showed that the intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  measured in air condition were significantly greater ( $p < 0.01$ ) than those recorded in the uterus, from CM *ex utero*, and from CM *in utero*, except that of  $F_0$  measured in the uterus ( $p > 0.05$ ). In the uterus condition, the intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were greater ( $p < 0.01$ ,  $p < 0.05$  for  $F_3$  CM *ex utero*) than that from CM *ex utero* and from CM *in utero*. In the CM conditions, only  $F_0$  ( $p < 0.05$ ) and  $F_1$  ( $p < 0.01$ ) from CM *ex utero* were significantly different from CM *in utero*.

For the female talker, the results indicated significant interaction between formant ( $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$ ) and recording location ( $F_{9,27} = 6.693$ ,  $p < 0.0001$ ), main effects for formant ( $F_{3,9} = 24.298$ ,  $p = 0.0001$ ) and location ( $F_{3,9} = 136.027$ ,  $p < 0.0001$ ). The post hoc multiple comparison test (Newman-Keuls) indicated that the intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  measured in air condition were significantly greater ( $p < 0.01$ ,  $p < 0.05$  for  $F_2$  in the uterus) than those recorded in the uterus, from CM *ex utero* and from CM *in utero*, but not that of  $F_0$  and  $F_1$  measured in the uterus ( $p > 0.05$ ). In the uterus recording condition, the intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were greater ( $p < 0.01$ ) than that from CM *ex utero* and from CM *in utero*. Between the conditions of CM *ex utero* and CM *in utero*, only  $F_1$  ( $p < 0.05$ ) and  $F_2$  ( $p < 0.01$ ) from CM *ex utero* were significantly different from CM *in utero*.

Table 4-28 contains the means and standard deviations of relative intensity levels of  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  recorded in different locations for the vowel /æ/ produced by the male and female talkers. The data are also graphically displayed in Figure 4-10. For the vowel /Λ/, the data are displayed in Table 4-31 and Figure 4-11. Similar to the vowel /ε/, vowels /æ/ and /Λ/ have high  $F_1$  frequencies, low  $F_2$  frequencies ( $< 2000$  Hz), and high  $F_3$  frequencies. The spectral analyses and statistical analyses (ANOVA) clearly showed the similarities of characteristics of transmission into the uterus and into the fetal inner ear *in utero* among the vowels /ε/, /æ/, and /Λ/. For both vowel /æ/ and /Λ/, in the uterus recordings,  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  were well received for both talkers. From CM *ex utero* recordings,  $F_0$ ,  $F_1$ , and  $F_2$  for both talkers were transmitted into the fetal inner ear, but  $F_3$  was close to the level of background noise. From CM *in utero* recordings,  $F_0$ ,  $F_1$ , and  $F_2$  of the vowel /æ/ were preserved for the female talker, but only  $F_0$  and  $F_1$  were received for the male talker, since  $F_2$  was close to the level of background noise. For the vowel /Λ/,  $F_0$ ,  $F_1$ , and  $F_2$  were preserved for both talkers,

Table 4-28. Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ ,  $F_3$ ) for vowel /æ/ produced by each talker at different recording locations in the 105 dB condition.

| Male talker, /æ/   |      | Condition |           |             |             |
|--------------------|------|-----------|-----------|-------------|-------------|
|                    |      | In Air    | In Uterus | CM-ex utero | CM-in utero |
| $F_0$              | Mean | 47.26     | 38.70     | 4.90        | 7.24        |
|                    | S.D. | 2.51      | 4.92      | 8.16        | 6.09        |
| $F_1$              | Mean | 48.26     | 37.64     | 21.62       | 13.98       |
|                    | S.D. | 6.45      | 5.52      | 10.05       | 8.43        |
| $F_2$              | Mean | 43.04     | 25.26     | 11.94       | 0.12        |
|                    | S.D. | 2.62      | 3.61      | 2.98        | 5.84        |
| $F_3$              | Mean | 33.52     | 6.72      | 0.90        | -0.34       |
|                    | S.D. | 5.65      | 4.99      | 3.95        | 2.10        |
| Female talker, /æ/ |      |           |           |             |             |
| $F_0$              | Mean | 42.78     | 33.40     | 11.68       | 17.96       |
|                    | S.D. | 3.56      | 3.65      | 1.99        | 1.62        |
| $F_1$              | Mean | 48.94     | 38.92     | 22.50       | 17.84       |
|                    | S.D. | 4.82      | 6.79      | 4.84        | 4.75        |
| $F_2$              | Mean | 48.70     | 33.88     | 15.30       | 5.56        |
|                    | S.D. | 4.74      | 6.07      | 7.04        | 0.99        |
| $F_3$              | Mean | 41.14     | 14.50     | 2.86        | -3.26       |
|                    | S.D. | 1.67      | 6.48      | 4.23        | 1.53        |

Figure 4-10. Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ , and  $F_3$ ) for vowel /æ/ produced by both talkers recorded at different locations at 105 dB SPL. Bars equal one standard deviation. Male talker results - upper panel; female talker results - lower panel.

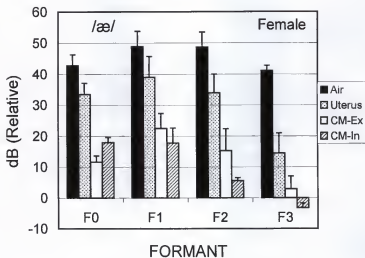
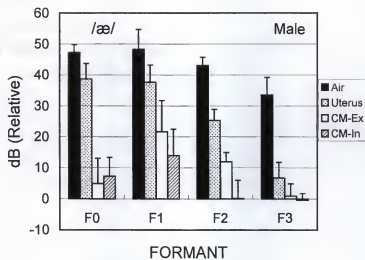
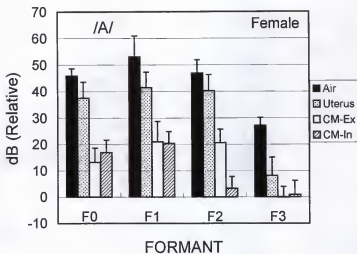
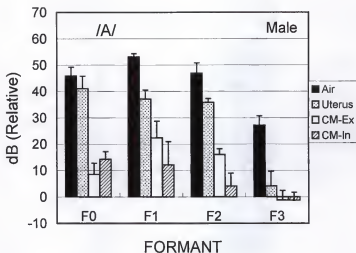


Table 4-29. Mean and standard deviation (S.D.) of relative intensity levels (dB) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ ,  $F_3$ ) for vowel /A/ produced by each talker at different recording locations in the 105 dB condition.

| Male talker, /A/   |      | Condition |           |             |             |
|--------------------|------|-----------|-----------|-------------|-------------|
|                    |      | In Air    | In Uterus | CM-ex utero | CM-in utero |
| $F_0$              | Mean | 45.90     | 41.12     | 8.48        | 14.18       |
|                    | S.D. | 3.29      | 4.65      | 4.24        | 2.91        |
| $F_1$              | Mean | 53.08     | 37.08     | 22.30       | 12.08       |
|                    | S.D. | 1.21      | 3.46      | 6.33        | 8.89        |
| $F_2$              | Mean | 46.90     | 35.80     | 16.02       | 4.04        |
|                    | S.D. | 3.92      | 1.45      | 2.14        | 4.85        |
| $F_3$              | Mean | 27.14     | 4.02      | -1.24       | -1.30       |
|                    | S.D. | 3.55      | 5.54      | 3.58        | 2.97        |
| Female talker, /A/ |      |           |           |             |             |
| $F_0$              | Mean | 44.22     | 37.44     | 13.24       | 17.00       |
|                    | S.D. | 2.67      | 6.10      | 5.39        | 4.64        |
| $F_1$              | Mean | 49.80     | 41.44     | 20.98       | 20.40       |
|                    | S.D. | 7.76      | 5.77      | 7.67        | 4.44        |
| $F_2$              | Mean | 46.96     | 40.22     | 20.56       | 3.20        |
|                    | S.D. | 5.04      | 6.01      | 5.12        | 4.56        |
| $F_3$              | Mean | 31.10     | 8.18      | -0.08       | 0.90        |
|                    | S.D. | 2.98      | 6.94      | 3.95        | 5.31        |

Figure 4-11. Mean of intensity levels (dB relative) of fundamental frequency ( $F_0$ ) and first three formant frequencies ( $F_1$ ,  $F_2$ , and  $F_3$ ) for vowel /A/ (=/Λ/) produced by both talkers recorded at different locations at 105 dB SPL. Bars equal one standard deviation. Male talker results - upper panel; female talker results - lower panel.





but  $F_2$  was only 5 dB above the level of background noise in the CM *in utero* recording condition. Thus, /æ/ and /ʌ/ could be well identified in the uterus recordings for both talkers, and also could be recognized from CM *ex utero* because  $F_2$  was well perceived (10-20 dB above the noise floor) for both talkers. Furthermore, the identification of the vowels /æ/ and /ʌ/ might be made from CM *in utero* recordings for both talkers, since  $F_2$  was still 5 dB above background noise, except of the vowel /æ/ by the male talker, and should be perceived in the fetal inner ear *in utero*. Table 4-30 provides a summary of information presented above regarding the characteristics of the five vowels.

The results from the spectrum analyses of vowels support the hypotheses that acoustic energy in the second and third formants measured in air for both male and female talkers will be reduced when recorded in the uterus, and will be reduced to the noise floor when recorded from fetal inner ear *in utero*.

The transmission of vowels into the uterus and into the fetal inner ear *in utero* follows the same pattern of low-pass filter characteristics as external sounds transmitted inside the uterus and to the fetal inner ear *in utero*. Low-frequency sounds penetrate the maternal tissues and fluids, and the fetal head more effectively than high frequencies (Gerhardt, Abrams and Oliver, 1990; Gerhardt et al., 1992). From air through the maternal tissues and fluids into the uterus, sounds are attenuated by 5-10 dB in the low-frequency range (< 1000 Hz) and 20-30 dB for higher frequencies (> 1000 Hz). To reach the fetal inner ear, the spectral contents of airborne sounds are further modified by the bone conduction route through the fetal head. For low frequencies from 125 to 250 Hz, airborne sounds would be reduced by 10-20 dB to reach the fetal inner ear. For frequencies from 500 to 2000 Hz, sounds would be reduced by 35-45 dB. In general, low-frequency

Table 4-30. Summary of acoustic analyses of vowels

| Vowel | Articulation    |      | Formant Frequency |                | Stimulus Present at F <sub>2</sub> |                    |
|-------|-----------------|------|-------------------|----------------|------------------------------------|--------------------|
|       | Tongue Position |      | F <sub>1</sub>    | F <sub>2</sub> | In Uterus                          | CM <i>in utero</i> |
| /i/   | Front           | High | Low               | High           | Yes                                | No                 |
| /I/   | Front           | High | Low               | High           | Yes                                | No                 |
| /e/   | Front           | Low  | High              | Low            | Yes                                | Yes                |
| /æ/   | Front           | Low  | High              | Low            | Yes                                | Yes                |
| /ʌ/   | Central         | Low  | High              | Low            | Yes                                | Yes                |

Note: Decisions about if the stimulus was present at F<sub>2</sub> (> 5 dB) were made from inspection of Figures 4-7 through 4-11.

components of external sounds are well perceived at the level of fetal inner ear, while high-frequency components are reduced to a great degree before reaching the fetal inner ear.

The data from the current study clearly showed that the fundamental frequency ( $F_0$ ) and the first three formants ( $F_1$ ,  $F_2$ , and  $F_3$ ) of all five vowels were well preserved in the uterus recordings for both the male and female talkers. These results are consistent with the high intelligibility scores obtained in this study. Querleu et al. (1988b) also found that  $F_2$  had critical effect on the recognition of French vowels recorded within the uterus.

The acoustic cues necessary for the perception of vowels lie in the patterns of formants. The first two lowest frequency formants are usually required to identify the vowels. Generally, two formants are required for front vowels, which have a high  $F_2$  frequency (/i/, /I/, /e/, and /æ/; /A/ is a central vowel). A single formant can be used to approximate the back vowels, which have a low  $F_2$  frequency (/u/, /U/, /o/, /O/, and /a/).  $F_3$  is more important for front vowels than for back vowels. However, the steady-state formant frequency patterns are not the only factors determining listener identification of vowels. For example, men, women, and children produce the same vowel with different formant frequencies. In this case, listeners must use general patterns for formant relationships rather than exact frequencies or even an exact ratio of frequencies. In addition, listeners also have to use contextual cues for vowel identification in those speakers who use a fast rate of speech (Borden and Harris, 1984).

In the present study, the spectral analyses of vowels from the CM recordings indicated that fundamental frequency ( $F_0$ ) and low-frequency formants,  $F_1$  and  $F_2$  (< 2000 Hz) were well preserved in the fetal inner ear *in utero*. For vowels /i/ and /I/ that have high frequency second formants (> 2000 Hz), only  $F_0$  and  $F_1$  were perceived in the recordings

from CM *in utero*. Whereas for vowels /e/, /æ/, and /ʌ/ that have low frequency second formants (<2000 Hz),  $F_0$ ,  $F_1$ , and  $F_2$  were all perceived in the recordings from CM *in utero*. Thus, vowels (/e/, /æ/, and /ʌ/) with low frequency second formants (<2000 Hz) might be easily identified from CM *in utero* recordings. Because  $F_3$  of all five vowels were higher than 2000 Hz ( $F_3$  of /i/ is even above 3000 Hz), the third formants were not preserved in the fetal inner ear *in utero*. The identification of vowels in the recordings from CM *in utero* for both male and female speakers might be easy because they are voiced, relative high in intensity, and have prominent formant frequencies.

## CHAPTER 5 SUMMARY AND CONCLUSIONS

This study had two distinct components. The first involved recording speech produced through a loudspeaker with an air microphone, a hydrophone placed in the uterus of a pregnant sheep, and an electrode surgically secured to the round window of the fetus *ex utero* and *in utero* (cochlear microphonic, CM). The speech stimuli consisted of two separate lists, Vowel-Consonant-Vowel (VCV) nonsense syllables and Consonant-Vowel-Consonant (CVC) monosyllable words spoken by a male and a female talker. They were presented at two airborne intensity levels, 105 and 95 dB SPL. Perceptual audio CDs were constructed from one recording with the best quality sound.

The second portion of the study involved playing the recordings to a group of normal hearing adults (N=139) over earphones. The intelligibility of speech was evaluated from the judges' responses to the speech stimuli under 16 different recording conditions.

The speech (VCV nonsense syllables and CVC words) intelligibility scores as a function of recording location alone, decreased from the air to the uterus locations and further decreased from the CM *ex utero* to the CM *in utero* conditions. Intelligibility was significantly higher for the recordings in air than in the uterus, and significantly higher for the recordings from CM *ex utero* than from CM *in utero*. In addition, the intelligibility scores of the male voice were significantly higher than that of the female

voice across all four recording locations for VCV nonsense syllables, but not for CVC words. The results also showed stimulus level effect on the intelligibility. Overall, when the mean intelligibility scores were averaged across two stimulus levels (105 and 95 dB SPL) and stimulus types (VCV and CVC stimuli), they were 91.2% and 84.7% for the male and female voices recorded within the uterus, respectively. Whereas, the mean intelligibility scores recorded from CM *in utero*, averaged across two levels and stimulus types for the male and female voices, were 45.0% and 41.6%, respectively. The recordings within the uterus reflect the speech energies present in the amniotic fluid, whereas the recordings from CM *in utero* represent the actual fetal physiological responses of the auditory periphery to externally generated speech.

Previous studies on the transmission of sound pressure into the maternal abdomen and uterus have shown consistent low-pass filter characteristics for external sound at the fetal head (Vince et al., 1982; Querleu et al., 1988a; Gerhardt, Abrams and Oliver, 1990; Richards et al., 1992; Peters et al., 1993a, 1993b). For frequencies less than 250 Hz, external sound passes through the uterus to the fetus with little reduction in sound pressure, and in some instances the pressure is greater within the uterus than it is outside the abdomen. Above 250 Hz, sound pressure attenuation occurs at a rate of approximately 6 dB per octave and reaches about 20 dB for 4000 Hz (Gerhardt, Abrams and Oliver, 1990). Thus, external speech signals would be shaped by the tissues and fluids of pregnancy before reaching the fetal head. Moreover, sound transmission properties through the fetal head to the inner ear by bone conduction further modified the stimulus (Gerhardt et al., 1992; Gerhardt et al., 1996). This influence coupled to the attenuation of sound pressures provided by the tissues and fluids of pregnancy result in

some isolation of the fetus from external sounds. Fetal sheep probably detect low-frequency sound produced outside its mother with a loss of 10 to 20 dB for 125 and 250 Hz, respectively. For frequencies from 500 to 2000 Hz, the fetus is isolated by 35-45 dB (Gerhardt et al., 1992). Therefore, the recordings of external speech from CM *in utero* would be degraded to a greater degree for the high-frequency components of speech rather than the low-frequency components. Intelligibility would be expected to follow.

The present findings showed much better intelligibility of speech recorded within the uterus than previously found (Querleu et al., 1988b; Griffiths et al., 1994). Querleu et al. (1988b) found that about 30% of 3120 French phonemes recorded within the uterus of pregnant women were recognized. Griffiths et al. (1994) showed the intelligibility of speech stimuli recorded within the uterus of a pregnant sheep, was 55% and 34% for the male and female talkers, respectively. However, from the current study the intelligibility was 91.2% and 84.7% for the male and female voices recorded in the uterus, respectively. The discrepancy might be accounted for by the higher stimulus levels and the use of earphones.

Consonant feature transmission was analyzed using SINFA. The results confirmed the findings that voicing information is well retained inside the uterus (Griffiths et al., 1994). Furthermore, the present study demonstrated that voicing information is also accurately represented in the fetal inner ear (CM recordings) *in utero*. Manner and place information were not maintained as well as voicing information at the fetal inner ear. These results are consistent with those of Miller and Nicely (1955), and Wang et al. (1978), in which low-pass filtering of speech signals resulted in a greater loss of manner and place information than of voicing information. They concluded that the

higher frequency information in the speech signal is critical for accurate identification of manner and place of articulation. Voicing information was well preserved in the fetal inner ear *in utero* after low-pass filtering by the tissues and fluids associated with pregnancy, and the fetal skull. However, manner and place information, high-frequency components of speech, were lost in the transmission to the fetal inner ear *in utero*, especially for the female voice.

In the present study, the results of spectral analyses of vowels clearly showed that the fundamental frequency ( $F_0$ ) and the first three formants ( $F_1$ ,  $F_2$ , and  $F_3$ ) of all five vowels (/i/, /I/, /e/, /æ/, /ʌ/) were well preserved in the uterus recordings for both the male and female talkers. They were also reflected in the results of high intelligibility scores obtained in this study. Querleu et al. (1988b) noted that  $F_2$  had a critical effect on the recognition of French vowels recorded within the uterus.

It is well known that the acoustic cues necessary for the identification of vowels lie in the patterns of the formants. The first two lowest frequency formants ( $F_1$  and  $F_2$ ) are usually required to identify the vowels. Generally, two formants are required for front vowels which have a high  $F_2$  frequency (/i/, /I/, /e/, and /æ/; /ʌ/ is a central vowel), a single formant ( $F_1$ ) can be used to approximate the back vowels which have a low  $F_2$  frequency (/u/, /U/, /o/, /O/, and /a/).  $F_3$  is more important for front vowels than for back vowels (Borden and Harris, 1984).

The data from the CM recordings indicated that fundamental frequency ( $F_0$ ) and low-frequency formants,  $F_1$  and  $F_2$  (< 2000 Hz) were well represented in the fetal inner ear *in utero*. For vowels /i/ and /I/ that have high-frequency  $F_2$  (> 2000 Hz), only  $F_0$  and  $F_1$  were perceived in recordings from CM *in utero*; whereas for vowels /e/, /æ/, and /ʌ/ that



have low-frequency  $F_2$  ( $< 2000$  Hz),  $F_0$ ,  $F_1$ , and  $F_2$  were all perceived in recording from CM *in utero*. Thus, vowels (/ε/, /æ/, and /Λ/) with low-frequency  $F_2$  ( $< 2000$  Hz) might be easily identified from CM *in utero* recordings. Because  $F_3$  of all five vowels were higher than 2000 Hz ( $F_3$  of /i/ is even above 3000 Hz),  $F_3$  were not preserved in the fetal inner ear *in utero*.

This study demonstrated that externally generated speech signals could reach the fetal inner ear *in utero*. The most relevant features of the speech signal for purposes of identification are received by the low-frequency content of the signal. Consistent with the low-pass filtering, by maternal tissues and fluids and fetal skull, of external generated sounds, voicing information is received by the fetal inner ear *in utero*, while speech energy conveying manner and place information is attenuated and less detected at the fetal inner ear. Male and female talker intelligibility scores averaged 45% and 42%, respectively, when recorded from the fetal CM *in utero*. They represent the speech energies received by the fetal inner ear *in utero*, which are underestimated by using round window electrode placements.

The implications of this research relate to theories regarding the prenatal functional development of auditory pathways and to the foundations for the later acquisition of speech and language (Cooper and Aslin, 1989; Querleu et al., 1989; Ruben, 1992; Abrams, Gerhardt and Antonelli, 1998). It has been postulated that prenatal sensory and learning experiences help to organize higher cortical function and provide the foundation for future learning abilities (Fifer and Moon, 1988; Hepper, 1992; Smotherman and Robinson, 1995). When discussing the concept of innate abilities, one should take into account the fact that a neonate is not without experience with speech stimuli.

# APPENDIX A

## SUBJECT RESPONSE SHEET

### VCV Nonsense Syllables

- |                    |                     |                    |                    |
|--------------------|---------------------|--------------------|--------------------|
| 1. /a <u>b</u> a/  | 2. /a <u>p</u> a/   | 3. /a <u>d</u> a/  | 4. /a <u>t</u> a/  |
| 5. /a <u>g</u> a/  | 6. /a <u>k</u> a/   | 7. /a <u>f</u> a/  | 8. /a <u>v</u> a/  |
| 9. /a <u>s</u> a/  | 10. /a <u>z</u> a/  | 11. /a <u>m</u> a/ | 12. /a <u>n</u> a/ |
| 13. /a <u>S</u> a/ | 14. /a <u>tS</u> a/ |                    |                    |

### CVC Words

- |  |  |   |  |
|--|--|---|--|
| 1. bass<br>batch<br>badge<br><u>bat</u><br>bash<br>back    | 2. loss<br><u>laws</u><br>lodge<br>log<br>long<br>lob  | 3. wick<br>with<br>wit<br><u>wig</u><br>witch<br>will   | 4. duff<br>duth<br><u>dumb</u><br>dove<br>dub<br>dug |
| 5. cup<br>cub<br>cud<br>come<br><u>cuff</u><br>cut         | 6. dim<br>did<br>dill<br>dip<br><u>dig</u><br>din      | 7. dung<br>duv<br>dug<br>dud<br><u>dun</u><br>dub       | 8. fit<br>fib<br>fig<br><u>fill</u><br>fin<br>fizz   |
| 9. leash<br><u>leave</u><br>liege<br>leach<br>lead<br>leap | 10. <u>toss</u><br>talks<br>tall<br>tog<br>tong<br>taj | 11. lag<br><u>lash</u><br>lath<br>lack<br>lass<br>laugh | 12. man<br><u>mat</u><br>mad<br>mack<br>mass<br>math |

- |     |   |     |   |     |   |     |  |
|-----|---|-----|---|-----|---|-----|--|
| 13. | base<br>bays<br>bayed<br><u>beige</u><br>bake<br>bathe        | 14. | pan<br><u>pass</u><br>pack<br>path<br>pad<br>pat    | 15. | peach<br>peas<br>peal<br>peat<br><u>peak</u><br>peace | 16. | pitch<br>pip<br>pig<br>pick<br>pill<br>pit           |
| 17. | pus<br>putt<br>puff<br>puck<br><u>pup</u><br>pub              | 18. | has<br>hag<br>have<br>half<br><u>hath</u><br>hash   | 19. | weave<br>wean<br>week<br>weed<br><u>we're</u><br>weal | 20. | sash<br>sack<br><u>sad</u><br>sap<br>sag<br>sat      |
| 21. | sheath<br>sheave<br>sheaf<br>sheik<br>sheathe<br><u>sheen</u> | 22. | sin<br>sill<br>sip<br>sick<br><u>sing</u><br>sit    | 23. | <u>sud</u><br>sup<br>sub<br>sum<br>sun<br>sung        | 24. | tam<br>tag<br>tap<br>tang<br>tan<br><u>tab</u>       |
| 25. | tear<br>teeth<br><u>teethe</u><br>teel<br>tease<br>team       | 26. | red<br>wed<br>dead<br><u>led</u><br>shed<br>fed     | 27. | <u>sold</u><br>hold<br>cold<br>told<br>gold<br>mold   | 28. | wig<br>rig<br>gig<br>big<br>pig<br><u>dig</u>        |
| 29. | thick<br>chick<br><u>kick</u><br>lick<br>sick<br>pick         | 30. | tin<br>kin<br><u>fin</u><br>shin<br>thin<br>pin     | 31. | mark<br>park<br>dark<br><u>bark</u><br>lark<br>shark  | 32. | tale<br><u>gale</u><br>male<br>bale<br>pale<br>rail  |
| 33. | feel<br>eel<br><u>peel</u><br>keel<br>reel<br>heel            | 34. | till<br>kill<br>hill<br>mill<br><u>will</u><br>bill | 35. | peal<br>zeal<br><u>feal</u><br>reel<br>veal<br>seal   | 36. | same<br>tame<br><u>shame</u><br>game<br>lame<br>came |

- |     |   |     |  |     |   |     |   |
|-----|---|-----|--|-----|---|-----|---|
| 37. | then<br><u>ten</u><br>fen<br>hen<br>den<br>pen      | 38. | fin<br>win<br><u>pin</u><br>din<br>sin<br>tin  | 39. | chin<br>gin<br>tin<br>sin<br>shin<br><u>thin</u>    | 40. | zee<br><u>thee</u><br>dee<br>knee<br>see<br>lee     |
| 41. | tent<br>pent<br>bent<br>dent<br><u>rent</u><br>went | 42. | rip<br>lip<br>chip<br>tip<br>dip<br><u>hip</u> | 43. | shop<br>pop<br><u>top</u><br>lop<br>cop<br>hop      | 44. | <u>yore</u><br>for<br>gore<br>wore<br>roar<br>lore  |
| 45. | fie<br>thy<br><u>vie</u><br>lie<br>thigh<br>high    | 46. | dip<br><u>zip</u><br>gyp<br>ship<br>nip<br>lip | 47. | <u>nest</u><br>west<br>best<br>rest<br>jest<br>vest | 48. | rust<br>gust<br><u>bust</u><br>lust<br>just<br>dust |
| 49. | rat<br><u>mat</u><br>bat<br>vat<br>fat<br>that      | 50. | may<br>they<br>gay<br>bay<br>nay<br><u>way</u> |     |   |     |   |

APPENDIX B  
RAW DATA FROM SUBJECT RESPONSE FORMS

The following tables contain the individual responses (number of correct responses) to VCV (A) and CVC (B) stimuli under 16 recording conditions.

Letter Codes:

A = In Air  
U = In Uterus  
X = CM-*ex utero*  
I = CM-*in utero*  
M = Male  
F = Female  
H = 105 dB  
L = 95 dB

## A. VCV Nonsense Syllables

| Subjects | Conditions |     |     |     |     |     |     |     |     |     |     |     |     |     | IFL |
|----------|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|          | AMH        | AML | UMH | UML | XMH | XML | IMH | IML | AFH | AFL | UPH | UFL | XFH | XFL |     |
| 1        | 14         | 14  | 12  | 14  | 12  | 11  | 5   | 4   | 14  | 13  | 12  | 11  | 8   | 8   | 7   |
| 2        | 14         | 14  | 13  | 14  | 12  | 10  | 7   | 4   | 13  | 13  | 10  | 9   | 6   | 7   | 5   |
| 3        | 14         | 14  | 12  | 13  | 11  | 9   | 7   | 3   | 13  | 13  | 12  | 12  | 6   | 5   | 3   |
| 4        | 14         | 14  | 13  | 14  | 13  | 11  | 5   | 6   | 13  | 13  | 13  | 12  | 8   | 7   | 5   |
| 5        | 13         | 13  | 14  | 13  | 11  | 10  | 8   | 2   | 13  | 13  | 11  | 13  | 5   | 9   | 5   |
| 6        | 14         | 14  | 13  | 14  | 12  | 9   | 5   | 3   | 13  | 13  | 11  | 10  | 8   | 6   | 4   |
| 7        | 14         | 14  | 10  | 13  | 10  | 10  | 6   | 6   | 13  | 13  | 13  | 10  | 6   | 5   | 4   |
| 8        | 14         | 13  | 13  | 14  | 11  | 8   | 6   | 6   | 13  | 13  | 13  | 12  | 7   | 6   | 5   |
| 9        | 14         | 14  | 14  | 14  | 10  | 12  | 6   | 5   | 12  | 12  | 10  | 12  | 6   | 6   | 2   |
| 10       | 14         | 14  | 13  | 14  | 13  | 12  | 8   | 5   | 12  | 12  | 11  | 11  | 7   | 6   | 4   |
| 11       | 14         | 14  | 13  | 14  | 13  | 11  | 7   | 6   | 12  | 13  | 13  | 12  | 8   | 10  | 6   |
| 12       | 14         | 14  | 13  | 14  | 8   | 10  | 7   | 4   | 13  | 12  | 11  | 12  | 8   | 7   | 4   |
| 13       | 14         | 14  | 12  | 14  | 10  | 10  | 6   | 6   | 13  | 13  | 12  | 13  | 7   | 8   | 6   |
| 14       | 14         | 14  | 12  | 14  | 11  | 11  | 8   | 5   | 13  | 14  | 13  | 11  | 7   | 7   | 3   |
| 15       | 14         | 14  | 13  | 14  | 13  | 11  | 4   | 3   | 12  | 13  | 14  | 11  | 5   | 9   | 4   |
| 16       | 13         | 13  | 13  | 13  | 9   | 9   | 7   | 5   | 13  | 13  | 9   | 7   | 5   | 6   | 5   |
| 17       | 14         | 14  | 14  | 14  | 13  | 11  | 11  | 5   | 12  | 12  | 12  | 9   | 6   | 4   | 3   |
| 18       | 13         | 14  | 12  | 14  | 12  | 8   | 5   | 7   | 13  | 13  | 13  | 10  | 7   | 8   | 3   |
| 19       | 14         | 14  | 11  | 13  | 10  | 8   | 7   | 3   | 13  | 12  | 11  | 10  | 7   | 7   | 6   |
| 20       | 14         | 13  | 12  | 13  | 11  | 8   | 7   | 6   | 10  | 12  | 10  | 10  | 6   | 6   | 2   |
| 21       | 14         | 13  | 13  | 14  | 10  | 10  | 5   | 6   | 12  | 13  | 10  | 11  | 8   | 6   | 4   |
| 22       | 14         | 14  | 11  | 12  | 13  | 9   | 5   | 2   | 13  | 12  | 12  | 13  | 6   | 5   | 3   |
| 23       | 14         | 14  | 11  | 12  | 9   | 8   | 3   | 4   | 12  | 12  | 10  | 11  | 7   | 5   | 8   |
| 24       | 14         | 14  | 13  | 12  | 11  | 8   | 4   | 5   | 13  | 13  | 12  | 11  | 9   | 5   | 4   |

Conditions

| Subjects | AMH | AML | UMH | UML | XMH | XML | IMH | IML | AFH | AFL | UFH | UFL | XFH | XFL | IFH | IFL |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 25       | 14  | 14  | 13  | 13  | 14  | 11  | 8   | 6   | 13  | 13  | 10  | 11  | 8   | 8   | 6   | 1   |
| 26       | 14  | 14  | 14  | 14  | 10  | 8   | 7   | 6   | 14  | 14  | 11  | 9   | 7   | 6   | 2   | 2   |
| 27       | 14  | 14  | 13  | 13  | 13  | 9   | 4   | 5   | 13  | 13  | 12  | 10  | 8   | 8   | 6   | 4   |
| 28       | 14  | 13  | 13  | 14  | 11  | 11  | 8   | 4   | 13  | 13  | 11  | 12  | 8   | 9   | 6   | 3   |
| 29       | 14  | 14  | 12  | 13  | 10  | 10  | 7   | 3   | 13  | 13  | 11  | 11  | 7   | 6   | 5   | 4   |
| 30       | 14  | 14  | 11  | 12  | 13  | 10  | 9   | 3   | 11  | 12  | 12  | 10  | 7   | 7   | 5   | 3   |
| 31       | 14  | 14  | 12  | 13  | 8   | 10  | 8   | 4   | 13  | 13  | 13  | 12  | 9   | 8   | 5   | 5   |
| 32       | 14  | 13  | 12  | 14  | 12  | 10  | 7   | 4   | 13  | 13  | 12  | 14  | 7   | 6   | 5   | 7   |
| 33       | 14  | 14  | 14  | 14  | 13  | 11  | 9   | 4   | 11  | 13  | 13  | 12  | 8   | 8   | 7   | 3   |

## B. CVC Words

| Test 1   |     | Conditions |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----------|-----|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Subjects | AMH | AML        | UMH | UML | XMH | XML | IMH | IML | AFH | AFL | UFH | UFL | XFH | XFL | IFH | IFL |
| 1        | 10  | 10         | 10  | 7   | 3   | 6   | 5   | 6   | 10  | 9   | 9   | 9   | 5   | 6   | 6   | 4   |
| 2        | 10  | 10         | 9   | 10  | 6   | 6   | 6   | 6   | 10  | 8   | 9   | 9   | 8   | 5   | 6   | 5   |
| 3        | 10  | 10         | 10  | 9   | 6   | 8   | 3   | 8   | 10  | 10  | 9   | 10  | 7   | 4   | 5   | 4   |
| 4        | 10  | 10         | 9   | 8   | 4   | 6   | 5   | 6   | 10  | 9   | 9   | 8   | 5   | 7   | 5   | 3   |
| 5        | 10  | 10         | 10  | 8   | 5   | 6   | 6   | 6   | 10  | 9   | 9   | 7   | 6   | 5   | 4   | 4   |
| 6        | 10  | 10         | 10  | 8   | 3   | 8   | 6   | 8   | 10  | 9   | 8   | 9   | 8   | 5   | 3   | 5   |
| 7        | 10  | 10         | 10  | 10  | 6   | 7   | 5   | 7   | 10  | 9   | 8   | 9   | 3   | 4   | 4   | 7   |
| 8        | 10  | 10         | 10  | 9   | 5   | 8   | 6   | 8   | 10  | 8   | 9   | 9   | 6   | 5   | 3   | 4   |
| 9        | 10  | 10         | 10  | 9   | 6   | 5   | 7   | 5   | 10  | 8   | 9   | 10  | 7   | 6   | 4   | 4   |
| 10       | 10  | 10         | 9   | 9   | 8   | 7   | 6   | 7   | 10  | 9   | 10  | 9   | 5   | 3   | 5   | 4   |
| 11       | 10  | 10         | 10  | 8   | 4   | 7   | 4   | 7   | 10  | 10  | 9   | 7   | 7   | 4   | 3   | 6   |
| 12       | 10  | 10         | 10  | 8   | 4   | 6   | 4   | 6   | 10  | 10  | 9   | 9   | 9   | 5   | 4   | 5   |
| 13       | 10  | 10         | 10  | 10  | 7   | 8   | 5   | 8   | 10  | 9   | 10  | 9   | 7   | 4   | 4   | 4   |
| 14       | 10  | 10         | 10  | 9   | 4   | 7   | 5   | 7   | 10  | 9   | 9   | 10  | 6   | 5   | 3   | 4   |
| 15       | 10  | 10         | 10  | 9   | 3   | 8   | 4   | 8   | 10  | 8   | 9   | 7   | 6   | 4   | 3   | 4   |
| 16       | 10  | 10         | 10  | 10  | 5   | 8   | 5   | 8   | 10  | 9   | 9   | 10  | 6   | 5   | 4   | 4   |
| 17       | 10  | 9          | 10  | 7   | 6   | 4   | 4   | 4   | 10  | 9   | 9   | 9   | 4   | 5   | 6   | 8   |
| 18       | 10  | 10         | 10  | 7   | 5   | 4   | 5   | 4   | 10  | 9   | 9   | 8   | 7   | 4   | 3   | 7   |
| 19       | 10  | 10         | 10  | 10  | 5   | 7   | 5   | 7   | 10  | 9   | 9   | 8   | 4   | 6   | 5   | 7   |



Test 2

| Subjects | Conditions |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----------|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|          | AMH        | AML | UMH | UML | XMH | XML | IMH | IML | AFH | AFL | UFH | UFL | XFH | XFL | IFH |
| 1        | 10         | 10  | 9   | 10  | 4   | 6   | 6   | 4   | 10  | 10  | 10  | 8   | 7   | 7   | 8   |
| 2        | 10         | 10  | 9   | 10  | 6   | 8   | 7   | 7   | 9   | 8   | 10  | 8   | 6   | 9   | 9   |
| 3        | 8          | 10  | 9   | 10  | 4   | 6   | 8   | 5   | 8   | 8   | 10  | 10  | 7   | 7   | 3   |
| 4        | 9          | 10  | 7   | 9   | 5   | 6   | 9   | 4   | 9   | 9   | 10  | 8   | 6   | 5   | 8   |
| 5        | 10         | 10  | 8   | 8   | 5   | 3   | 7   | 5   | 10  | 9   | 10  | 9   | 7   | 7   | 5   |
| 6        | 10         | 10  | 8   | 10  | 6   | 5   | 8   | 3   | 10  | 9   | 10  | 9   | 6   | 10  | 4   |
| 7        | 10         | 10  | 8   | 9   | 5   | 7   | 7   | 6   | 9   | 9   | 10  | 9   | 7   | 9   | 4   |
| 8        | 9          | 10  | 8   | 8   | 5   | 3   | 6   | 5   | 10  | 9   | 10  | 9   | 6   | 7   | 5   |
| 9        | 8          | 10  | 9   | 10  | 3   | 3   | 7   | 3   | 9   | 8   | 9   | 9   | 4   | 7   | 6   |
| 10       | 10         | 10  | 9   | 8   | 5   | 5   | 7   | 2   | 10  | 9   | 9   | 8   | 7   | 6   | 2   |
| 11       | 10         | 10  | 9   | 8   | 6   | 7   | 8   | 3   | 10  | 10  | 10  | 8   | 7   | 8   | 4   |
| 12       | 10         | 10  | 8   | 10  | 5   | 5   | 8   | 5   | 8   | 8   | 10  | 10  | 8   | 8   | 7   |
| 13       | 8          | 10  | 9   | 10  | 5   | 6   | 7   | 4   | 9   | 9   | 10  | 7   | 8   | 6   | 7   |
| 14       | 9          | 10  | 8   | 10  | 6   | 6   | 8   | 6   | 9   | 9   | 9   | 10  | 9   | 10  | 7   |
| 15       | 10         | 10  | 9   | 10  | 4   | 6   | 4   | 2   | 10  | 9   | 10  | 8   | 9   | 9   | 6   |
| 16       | 10         | 10  | 9   | 10  | 5   | 4   | 8   | 5   | 9   | 10  | 9   | 9   | 6   | 6   | 4   |
| 17       | 10         | 10  | 8   | 9   | 6   | 5   | 7   | 2   | 10  | 8   | 10  | 10  | 6   | 5   | 4   |
| 18       | 10         | 10  | 9   | 8   | 5   | 7   | 7   | 3   | 10  | 10  | 10  | 7   | 6   | 6   | 2   |
| 19       | 10         | 10  | 8   | 9   | 5   | 4   | 6   | 6   | 10  | 10  | 10  | 9   | 8   | 5   | 4   |
| 20       | 10         | 10  | 8   | 9   | 6   | 5   | 5   | 4   | 10  | 9   | 10  | 8   | 5   | 9   | 7   |
| 21       | 10         | 10  | 9   | 8   | 4   | 3   | 7   | 2   | 10  | 10  | 10  | 9   | 7   | 7   | 5   |

Test 3

| Subjects | Conditions |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----------|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|          | AMH        | AML | UMH | UML | XMH | XML | IMH | IML | AFH | AFL | UFH | UFL | XFH | XPL | IFH | IFL |
| 1        | 9          | 10  | 10  | 7   | 2   | 4   | 7   | 3   | 10  | 10  | 8   | 6   | 5   | 8   | 6   | 3   |
| 2        | 10         | 9   | 7   | 7   | 5   | 3   | 6   | 5   | 10  | 10  | 9   | 10  | 5   | 5   | 7   | 7   |
| 3        | 10         | 10  | 10  | 7   | 8   | 3   | 7   | 6   | 10  | 10  | 9   | 9   | 7   | 7   | 4   | 7   |
| 4        | 10         | 10  | 10  | 5   | 6   | 5   | 5   | 3   | 10  | 10  | 10  | 10  | 3   | 7   | 5   | 4   |
| 5        | 10         | 10  | 10  | 7   | 6   | 6   | 5   | 6   | 10  | 10  | 9   | 9   | 3   | 6   | 4   | 7   |
| 6        | 9          | 10  | 9   | 7   | 8   | 4   | 9   | 5   | 10  | 10  | 9   | 10  | 7   | 6   | 8   | 3   |
| 7        | 10         | 9   | 10  | 7   | 4   | 3   | 5   | 7   | 9   | 10  | 9   | 10  | 3   | 9   | 5   | 6   |
| 8        | 10         | 10  | 10  | 6   | 4   | 5   | 5   | 6   | 10  | 10  | 9   | 10  | 3   | 7   | 5   | 7   |
| 9        | 10         | 10  | 10  | 7   | 6   | 4   | 5   | 3   | 10  | 10  | 9   | 10  | 4   | 7   | 7   | 5   |
| 10       | 10         | 9   | 9   | 7   | 5   | 5   | 6   | 4   | 10  | 9   | 9   | 10  | 4   | 8   | 5   | 7   |
| 11       | 9          | 10  | 10  | 6   | 5   | 4   | 5   | 6   | 10  | 10  | 9   | 10  | 4   | 7   | 5   | 5   |
| 12       | 9          | 9   | 8   | 7   | 5   | 3   | 6   | 5   | 9   | 10  | 9   | 10  | 5   | 7   | 4   | 5   |
| 13       | 10         | 9   | 10  | 7   | 7   | 8   | 6   | 6   | 10  | 10  | 9   | 10  | 6   | 8   | 5   | 2   |
| 14       | 10         | 9   | 9   | 6   | 5   | 3   | 5   | 4   | 10  | 8   | 9   | 9   | 4   | 5   | 6   | 5   |
| 15       | 10         | 10  | 10  | 6   | 5   | 6   | 5   | 4   | 10  | 10  | 9   | 10  | 5   | 5   | 5   | 4   |
| 16       | 10         | 9   | 9   | 7   | 6   | 6   | 5   | 3   | 10  | 10  | 9   | 8   | 3   | 9   | 5   | 5   |
| 17       | 10         | 9   | 9   | 7   | 7   | 4   | 5   | 5   | 10  | 10  | 9   | 9   | 3   | 6   | 4   | 6   |
| 18       | 9          | 10  | 10  | 7   | 8   | 4   | 7   | 5   | 10  | 9   | 9   | 9   | 5   | 7   | 3   | 6   |
| 19       | 10         | 10  | 9   | 6   | 7   | 5   | 5   | 6   | 10  | 10  | 8   | 9   | 7   | 6   | 6   | 7   |
| 20       | 10         | 10  | 10  | 7   | 6   | 6   | 8   | 4   | 9   | 10  | 9   | 10  | 5   | 6   | 4   | 7   |

| Test 4   |     | Conditions |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|----------|-----|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Subjects | AMH | AML        | UMH | UML | XMH | XML | IMH | IML | APH | AFL | UFH | UFL | XFH | XFL | IFH | IFL |  |
| 1        | 9   | 9          | 8   | 10  | 8   | 5   | 5   | 3   | 10  | 9   | 9   | 10  | 6   | 4   | 5   | 5   |  |
| 2        | 10  | 10         | 10  | 7   | 6   | 5   | 6   | 2   | 10  | 10  | 8   | 8   | 8   | 8   | 6   | 4   |  |
| 3        | 10  | 10         | 9   | 7   | 8   | 6   | 5   | 4   | 10  | 10  | 9   | 9   | 6   | 6   | 6   | 4   |  |
| 4        | 9   | 10         | 10  | 10  | 6   | 4   | 5   | 3   | 10  | 10  | 9   | 10  | 4   | 7   | 5   | 5   |  |
| 5        | 10  | 9          | 9   | 9   | 7   | 5   | 6   | 4   | 10  | 9   | 9   | 10  | 7   | 4   | 3   | 7   |  |
| 6        | 10  | 10         | 9   | 7   | 8   | 5   | 6   | 2   | 10  | 10  | 9   | 9   | 9   | 6   | 5   | 4   |  |
| 7        | 10  | 9          | 9   | 9   | 7   | 8   | 6   | 3   | 10  | 10  | 10  | 10  | 9   | 8   | 3   | 6   |  |
| 8        | 10  | 10         | 10  | 8   | 8   | 5   | 7   | 2   | 10  | 10  | 9   | 9   | 6   | 8   | 5   | 4   |  |
| 9        | 10  | 10         | 9   | 9   | 8   | 6   | 6   | 2   | 10  | 10  | 9   | 9   | 9   | 6   | 6   | 4   |  |
| 10       | 10  | 10         | 10  | 7   | 8   | 7   | 7   | 4   | 10  | 9   | 8   | 9   | 7   | 7   | 5   | 4   |  |
| 11       | 9   | 10         | 9   | 10  | 6   | 7   | 5   | 4   | 10  | 10  | 10  | 9   | 6   | 5   | 7   | 4   |  |
| 12       | 10  | 10         | 10  | 8   | 8   | 6   | 5   | 2   | 10  | 10  | 9   | 10  | 9   | 8   | 6   | 3   |  |
| 13       | 10  | 9          | 10  | 8   | 9   | 6   | 6   | 4   | 10  | 10  | 8   | 9   | 7   | 6   | 4   | 3   |  |
| 14       | 10  | 10         | 10  | 10  | 7   | 5   | 7   | 5   | 9   | 10  | 9   | 10  | 6   | 7   | 5   | 2   |  |
| 15       | 10  | 9          | 9   | 10  | 6   | 6   | 7   | 1   | 10  | 10  | 10  | 9   | 8   | 9   | 4   | 7   |  |
| 16       | 10  | 9          | 9   | 9   | 9   | 6   | 7   | 4   | 10  | 9   | 10  | 10  | 9   | 10  | 4   | 3   |  |
| 17       | 9   | 9          | 9   | 7   | 9   | 5   | 5   | 5   | 10  | 9   | 9   | 10  | 9   | 6   | 4   | 4   |  |
| 18       | 9   | 9          | 9   | 8   | 9   | 5   | 6   | 5   | 10  | 9   | 9   | 10  | 4   | 6   | 4   | 5   |  |
| 19       | 10  | 10         | 10  | 9   | 7   | 4   | 7   | 4   | 10  | 9   | 10  | 10  | 7   | 8   | 5   | 6   |  |
| 20       | 10  | 10         | 10  | 8   | 8   | 7   | 6   | 3   | 10  | 9   | 10  | 10  | 7   | 7   | 6   | 7   |  |
| 21       | 10  | 10         | 10  | 8   | 8   | 7   | 5   | 3   | 10  | 10  | 9   | 10  | 5   | 7   | 5   | 4   |  |

Test 5

| Test 5   |     | Conditions |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----------|-----|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Subjects | AMH | AML        | UMH | UML | XMH | XML | IMH | IML | AFH | AFL | UFH | UFL | XFH | XFL | IFH | IFL |
| 1        | 8   | 10         | 10  | 10  | 9   | 4   | 8   | 5   | 10  | 9   | 8   | 9   | 7   | 4   | 2   | 4   |
| 2        | 9   | 9          | 9   | 9   | 7   | 7   | 5   | 5   | 10  | 9   | 10  | 7   | 8   | 3   | 7   | 3   |
| 3        | 9   | 10         | 9   | 10  | 9   | 8   | 7   | 7   | 10  | 9   | 9   | 7   | 7   | 4   | 7   | 4   |
| 4        | 8   | 10         | 10  | 10  | 5   | 6   | 5   | 4   | 10  | 9   | 8   | 6   | 7   | 6   | 3   | 4   |
| 5        | 9   | 10         | 10  | 9   | 6   | 5   | 4   | 4   | 10  | 9   | 8   | 7   | 7   | 6   | 5   | 5   |
| 6        | 9   | 9          | 9   | 10  | 7   | 7   | 7   | 5   | 10  | 9   | 8   | 8   | 6   | 5   | 7   | 5   |
| 7        | 9   | 10         | 10  | 10  | 8   | 6   | 7   | 6   | 10  | 9   | 8   | 8   | 10  | 7   | 6   | 6   |
| 8        | 8   | 9          | 9   | 10  | 7   | 7   | 6   | 8   | 10  | 9   | 10  | 5   | 7   | 5   | 8   | 5   |
| 9        | 9   | 10         | 10  | 10  | 8   | 6   | 7   | 6   | 9   | 9   | 8   | 7   | 6   | 6   | 7   | 5   |
| 10       | 9   | 10         | 10  | 10  | 7   | 5   | 5   | 5   | 9   | 9   | 8   | 8   | 7   | 8   | 9   | 5   |
| 11       | 9   | 10         | 10  | 10  | 8   | 6   | 5   | 6   | 10  | 9   | 8   | 7   | 5   | 6   | 3   | 4   |
| 12       | 9   | 9          | 9   | 10  | 7   | 8   | 5   | 5   | 10  | 9   | 9   | 8   | 8   | 7   | 5   | 3   |
| 13       | 9   | 10         | 10  | 10  | 9   | 6   | 8   | 5   | 10  | 9   | 8   | 7   | 7   | 6   | 5   | 4   |
| 14       | 9   | 9          | 9   | 9   | 7   | 8   | 5   | 7   | 10  | 9   | 10  | 7   | 7   | 9   | 6   | 6   |
| 15       | 9   | 10         | 10  | 10  | 9   | 5   | 8   | 6   | 10  | 9   | 9   | 8   | 6   | 5   | 4   | 4   |
| 16       | 9   | 9          | 10  | 8   | 5   | 5   | 4   | 4   | 8   | 9   | 7   | 7   | 6   | 5   | 7   | 2   |
| 17       | 8   | 10         | 10  | 10  | 6   | 7   | 7   | 6   | 10  | 9   | 8   | 7   | 6   | 6   | 7   | 6   |
| 18       | 9   | 9          | 9   | 9   | 7   | 8   | 6   | 3   | 10  | 9   | 10  | 8   | 5   | 5   | 7   | 3   |
| 19       | 9   | 9          | 9   | 9   | 7   | 8   | 6   | 3   | 10  | 9   | 8   | 8   | 7   | 5   | 6   | 4   |
| 20       | 9   | 9          | 10  | 10  | 7   | 9   | 5   | 4   | 10  | 9   | 9   | 8   | 8   | 6   | 8   | 4   |
| 21       | 9   | 9          | 9   | 10  | 10  | 5   | 6   | 6   | 10  | 9   | 8   | 7   | 6   | 5   | 8   | 2   |
| 22       | 9   | 10         | 10  | 9   | 8   | 7   | 8   | 5   | 9   | 9   | 8   | 7   | 6   | 3   | 6   | 3   |
| 23       | 9   | 10         | 8   | 9   | 9   | 7   | 6   | 5   | 10  | 9   | 8   | 8   | 9   | 5   | 3   | 3   |
| 24       | 9   | 10         | 9   | 10  | 7   | 9   | 7   | 5   | 10  | 9   | 9   | 7   | 6   | 8   | 6   | 6   |
| 25       | 9   | 10         | 10  | 9   | 6   | 7   | 5   | 6   | 9   | 9   | 8   | 7   | 7   | 6   | 7   | 3   |

APPENDIX C  
RAW DATA FROM ACOUSTIC ANALYSES OF VOWELS

The following tables contain the values of spectral analyses of vowels under different recording locations for male and female speakers. A. In air; B. In the uterus; C. CM-*ex utero*; D. CM-*in utero*.



| Female | Vowel | Word | Formant (Hz)   |                |                | Stimulus Level (dB) |                |                | Noise Level (dB) |                |                | Relative dB    |                |                |
|--------|-------|------|----------------|----------------|----------------|---------------------|----------------|----------------|------------------|----------------|----------------|----------------|----------------|----------------|
|        |       |      | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>0</sub>      | F <sub>1</sub> | F <sub>2</sub> | F <sub>0</sub>   | F <sub>1</sub> | F <sub>2</sub> | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> |
| /i/    | leave | 235  | 352            | 2809           | 3285           | 14.6                | 45.8           | 38.2           | 51.3             | 65.4           | 59.6           | 77.2           | 83.1           | 39.0           |
|        |       | 235  | 358            | 2763           | 3423           | 17.9                | 42.7           | 37.0           | 46.2             | 56.2           | 64.1           | 81.2           | 79.9           | 38.3           |
|        |       | 245  | 364            | 2951           | 3452           | 5.8                 | 46.0           | 34.4           | 51.6             | 59.8           | 58.3           | 79.2           | 84.4           | 54.0           |
|        |       | 245  | 357            | 2974           | 3463           | 10.5                | 45.6           | 38.8           | 51.5             | 51.7           | 62.0           | 80.3           | 79.1           | 41.2           |
|        |       | 248  | 333            | 2708           | 3060           | 13.2                | 41.4           | 41.0           | 54.1             | 63.0           | 67.5           | 71.1           | 80.9           | 49.8           |
| /l/    | wig   | 218  | 433            | 2382           | 3025           | 16.1                | 10.7           | 41.6           | 41.2             | 62.1           | 62.9           | 76.3           | 73.7           | 46.0           |
|        |       | 216  | 429            | 2362           | 2786           | 16.5                | 10.1           | 34.5           | 37.9             | 59.5           | 60.4           | 82.9           | 79.2           | 43.0           |
|        |       | 216  | 429            | 2362           | 2786           | 19.2                | 16.3           | 58.2           | 54.8             | 57.3           | 62.8           | 80.6           | 75.5           | 38.1           |
|        |       | 234  | 456            | 2219           | 2838           | 19.7                | 21.0           | 36.0           | 40.0             | 67.0           | 67.3           | 78.2           | 84.6           | 47.3           |
|        |       | 234  | 462            | 2554           | 3009           | 10.1                | 17.0           | 36.0           | 43.9             | 58.9           | 70.9           | 77.1           | 76.8           | 48.8           |
| /ε/    | led   | 229  | 670            | 2020           | 2903           | 15.1                | 17.6           | 27.7           | 36.4             | 61.1           | 74.4           | 81.1           | 82.1           | 46.0           |
|        |       | 226  | 675            | 2037           | 2951           | 13.5                | 21.6           | 38.7           | 49.1             | 55.0           | 71.2           | 77.2           | 76.1           | 41.5           |
|        |       | 218  | 622            | 1891           | 2786           | 16.4                | 26.4           | 28.5           | 37.2             | 63.2           | 71.2           | 75.1           | 79.5           | 46.8           |
|        |       | 208  | 634            | 1951           | 2740           | 14.3                | 25.5           | 34.6           | 43.0             | 53.0           | 65.4           | 76.8           | 80.8           | 38.7           |
|        |       | 216  | 1061           | 1883           | 2939           | 20.3                | 29.0           | 32.0           | 37.4             | 66.0           | 77.4           | 79.0           | 80.3           | 45.7           |
| /æ/    | bat   | 218  | 1053           | 1907           | 2951           | 19.5                | 19.6           | 21.3           | 36.3             | 64.5           | 73.6           | 78.4           | 77.0           | 45.0           |
|        |       | 216  | 1053           | 1907           | 2974           | 19.6                | 29.3           | 31.3           | 40.4             | 58.8           | 72.1           | 77.4           | 78.9           | 39.2           |
|        |       | 216  | 1079           | 1938           | 2786           | 18.6                | 27.4           | 31.5           | 38.6             | 64.0           | 80.9           | 78.9           | 80.4           | 45.4           |
|        |       | 221  | 1098           | 1997           | 2871           | 19.5                | 33.1           | 32.0           | 37.7             | 58.1           | 79.1           | 77.9           | 79.5           | 38.6           |
|        |       | 208  | 627            | 1252           | 2695           | 16.2                | 23.4           | 26.5           | 49.9             | 62.7           | 72.0           | 74.8           | 77.0           | 46.5           |
| /ʌ/    | dumb  | 216  | 704            | 1292           | 2809           | 22.6                | 29.9           | 35.9           | 44.3             | 64.0           | 83.8           | 90.4           | 77.4           | 41.4           |
|        |       | 216  | 643            | 1273           | 2763           | 13.4                | 24.4           | 32.5           | 45.4             | 60.4           | 72.0           | 77.3           | 78.5           | 47.0           |
|        |       | 242  | 712            | 1250           | 2838           | 18.6                | 25.3           | 33.2           | 48.7             | 60.1           | 78.0           | 79.6           | 82.2           | 41.5           |
|        |       | 221  | 664            | 1325           | 2838           | 16.2                | 21.3           | 35.5           | 46.3             | 60.9           | 74.1           | 76.3           | 75.0           | 44.7           |
|        |       | 221  | 664            | 1325           | 2838           | 16.2                | 21.3           | 35.5           | 46.3             | 60.9           | 74.1           | 76.3           | 75.0           | 44.7           |

## B. In Uterus

| Male | Vowel | Word   | Formant (Hz)   |                |                | Stimulus Level (dB) |                |                | Noise Level (dB) |                |                | Relative dB    |                |                |
|------|-------|--------|----------------|----------------|----------------|---------------------|----------------|----------------|------------------|----------------|----------------|----------------|----------------|----------------|
|      |       |        | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub>      | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub>   | F <sub>3</sub> | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub> |
| /i/  | leave | peak   | 132            | 353            | 2437           | 3630                | 14.6           | 53.3           | 53.5             | 66.8           | 52.6           | 70.8           | 74.9           | 73.8           |
|      |       | sheen  | 124            | 361            | 2409           | 3587                | 16.6           | 33.8           | 63.4             | 83.9           | 52.3           | 71.2           | 75.9           | 72.3           |
|      |       | teethe | 135            | 357            | 2584           | 3504                | 10.4           | 54.7           | 69.0             | 67.8           | 55.1           | 74.1           | 69.6           | 79.6           |
|      |       | peel   | 126            | 337            | 2524           | 3546                | 26.3           | 34.6           | 66.7             | 73.0           | 56.5           | 70.3           | 71.8           | 75.2           |
| /l/  | wig   | dig    | 163            | 317            | 2495           | 3716                | 13.9           | 35.0           | 66.6             | 71.4           | 62.8           | 71.5           | 74.5           | 74.6           |
|      |       | fill   | 116            | 374            | 2117           | 2614                | 12.0           | 17.1           | 66.9             | 73.3           | 51.4           | 67.1           | 73.6           | 72.6           |
|      |       | pick   | 113            | 397            | 2093           | 2740                | 14.1           | 15.0           | 51.7             | 56.6           | 54.7           | 66.8           | 75.3           | 75.6           |
|      |       | sing   | 124            | 406            | 1928           | 2614                | 11.0           | 5.3            | 45.5             | 66.5           | 59.4           | 62.6           | 77.9           | 72.4           |
| /e/  | led   | ten    | 127            | 357            | 1997           | 2645                | 15.6           | 34.4           | 60.0             | 72.0           | 61.7           | 70.0           | 73.7           | 77.1           |
|      |       | rent   | 135            | 451            | 2245           | 2838                | 8.2            | 17.5           | 52.7             | 53.3           | 61.2           | 67.6           | 73.3           | 78.0           |
|      |       | nest   | 117            | 626            | 1735           | 2805                | 12.2           | 34.1           | 48.9             | 65.1           | 55.4           | 69.8           | 76.0           | 72.9           |
|      |       | but    | 135            | 591            | 1928           | 2772                | 11.2           | 32.6           | 49.8             | 61.9           | 57.7           | 70.6           | 74.3           | 75.3           |
| /æ/  | lash  | mat    | 124            | 538            | 1755           | 2676                | 10.6           | 23.2           | 49.5             | 63.8           | 50.3           | 68.1           | 72.1           | 71.7           |
|      |       | pass   | 138            | 582            | 1905           | 2809                | 7.6            | 34.5           | 53.5             | 68.1           | 56.6           | 77.0           | 72.9           | 76.5           |
|      |       | hath   | 112            | 688            | 1776           | 2437                | 10.7           | 28.5           | 51.9             | 60.5           | 54.8           | 74.3           | 78.6           | 74.9           |
|      |       | dumb   | 113            | 696            | 1655           | 2614                | 11.7           | 32.8           | 49.9             | 71.0           | 53.8           | 71.5           | 73.7           | 74.1           |
| /ʌ/  | cuff  | pup    | 115            | 626            | 1735           | 2645                | 14.1           | 33.5           | 46.4             | 73.3           | 53.2           | 71.9           | 76.5           | 74.9           |
|      |       | sud    | 113            | 672            | 1598           | 2614                | 23.0           | 39.4           | 49.8             | 71.7           | 54.5           | 70.7           | 75.2           | 78.3           |
|      |       | hath   | 113            | 649            | 1617           | 2524                | 16.8           | 38.7           | 51.3             | 66.3           | 53.5           | 72.7           | 71.6           | 74.2           |
|      |       | peel   | 116            | 634            | 1192           | 2772                | 8.9            | 34.3           | 38.9             | 69.3           | 52.5           | 72.7           | 75.0           | 80.9           |
|      | dumb  | peel   | 113            | 641            | 1206           | 2645                | 15.5           | 35.1           | 36.3             | 72.5           | 51.1           | 69.9           | 72.1           | 76.8           |
|      |       | peel   | 113            | 626            | 1235           | 2740                | 11.9           | 31.4           | 37.1             | 64.6           | 56.1           | 74.0           | 73.5           | 71.4           |
|      |       | peel   | 113            | 605            | 1235           | 2645                | 19.4           | 35.1           | 37.3             | 72.5           | 56.0           | 70.2           | 74.6           | 71.9           |
|      |       | peel   | 135            | 605            | 1341           | 2676                | 13.2           | 35.3           | 47.1             | 73.7           | 58.8           | 69.8           | 80.5           | 71.7           |





## C. CM-ex utero

| Male | Vowel | Word   | Formant (Hz)   |                |                | Stimulus Level (dB) |                |                |                |                | Noise Level (dB) |                |                |                |                | Relative dB    |                |                |
|------|-------|--------|----------------|----------------|----------------|---------------------|----------------|----------------|----------------|----------------|------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
|      |       |        | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub>      | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub> | F <sub>0</sub>   | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub> | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub> |
| /i/  |       | leave  | 132            | 353            | 2437           | 3630                | 52.2           | 43.7           | 50.6           | 51.9           | 55.5             | 52.9           | 52.2           | 53.3           | 3.3            | 9.2            | 1.6            | 1.4            |
|      |       | peak   | 124            | 361            | 2409           | 3587                | 49.1           | 27.6           | 51.2           | 49.6           | 51.3             | 52.8           | 49.9           | 53.5           | 2.2            | 25.2           | -1.3           | 3.9            |
|      |       | sheen  | 135            | 357            | 2584           | 3504                | 43.4           | 32.8           | 50.1           | 50.3           | 54.4             | 48.1           | 52.8           | 51.7           | 11.0           | 15.3           | 2.7            | 1.4            |
|      |       | teethe | 126            | 337            | 2524           | 3546                | 37.2           | 22.2           | 44.3           | 39.2           | 43.2             | 45.8           | 49.8           | 42.2           | 6.0            | 23.6           | 5.5            | 3.0            |
|      |       | peel   | 163            | 317            | 2495           | 3716                | 44.5           | 26.0           | 53.2           | 53.4           | 50.5             | 51.6           | 50.7           | 52.2           | 6.0            | 25.6           | -2.5           | -1.2           |
| /ɪ/  |       | wig    | 116            | 374            | 2117           | 2614                | 48.2           | 16.5           | 53.6           | 51.9           | 50.0             | 50.2           | 51.8           | 49.0           | 1.8            | 33.7           | -1.8           | -2.9           |
|      |       | dig    | 113            | 397            | 2093           | 2740                | 44.9           | 15.8           | 49.1           | 46.8           | 52.2             | 50.2           | 50.0           | 63.3           | 7.3            | 34.4           | 0.9            | 16.5           |
|      |       | fill   | 124            | 406            | 1928           | 2614                | 42.7           | 10.4           | 44.3           | 53.3           | 52.6             | 47.7           | 49.2           | 48.3           | 9.9            | 37.3           | 4.9            | -5.0           |
|      |       | pick   | 127            | 357            | 1997           | 2645                | 48.5           | 22.5           | 48.9           | 47.4           | 53.3             | 48.3           | 47.3           | 56.9           | 4.8            | 25.8           | -1.6           | 9.5            |
|      |       | sing   | 135            | 451            | 2245           | 2838                | 45.8           | 16.1           | 47.6           | 52.3           | 52.9             | 50.5           | 47.6           | 52.6           | 7.1            | 34.4           | 0.0            | 0.3            |
| /e/  |       | led    | 117            | 626            | 1735           | 2805                | 46.6           | 31.7           | 49.0           | 51.6           | 48.8             | 48.9           | 55.2           | 52.4           | 2.2            | 17.2           | 6.2            | 0.8            |
|      |       | ten    | 135            | 591            | 1928           | 2772                | 46.4           | 22.7           | 46.1           | 50.9           | 53.2             | 55.3           | 48.1           | 51.7           | 6.8            | 32.6           | 2.0            | 0.8            |
|      |       | rent   | 124            | 538            | 1755           | 2676                | 43.8           | 15.4           | 47.5           | 51.1           | 55.8             | 49.3           | 50.0           | 51.3           | 12.0           | 33.9           | 2.5            | 0.2            |
|      |       | nest   | 138            | 582            | 1905           | 2809                | 38.5           | 23.0           | 45.6           | 54.9           | 53.4             | 51.9           | 58.2           | 49.3           | 14.9           | 28.9           | 12.6           | -5.6           |
|      |       | bat    | 112            | 688            | 1776           | 2437                | 44.1           | 21.5           | 44.7           | 49.1           | 49.7             | 54.0           | 52.5           | 44.3           | 5.6            | 32.5           | 7.8            | -4.8           |
| /æ/  |       | lash   | 113            | 696            | 1655           | 2614                | 41.3           | 23.5           | 36.4           | 48.3           | 47.8             | 51.8           | 50.9           | 49.5           | 6.5            | 28.3           | 14.5           | 1.2            |
|      |       | mat    | 115            | 626            | 1735           | 2645                | 42.8           | 28.0           | 38.8           | 50.7           | 57.0             | 53.2           | 51.1           | 52.9           | 14.2           | 25.2           | 12.3           | 2.2            |
|      |       | pass   | 113            | 672            | 1598           | 2614                | 56.4           | 32.5           | 39.7           | 51.6           | 48.1             | 45.1           | 54.6           | 51.4           | -8.3           | 12.6           | 14.9           | -0.2           |
|      |       | hath   | 113            | 649            | 1617           | 2524                | 44.5           | 39.9           | 39.6           | 49.1           | 51.0             | 49.4           | 49.8           | 55.2           | 6.5            | 9.5            | 10.2           | 6.1            |
|      |       | dumb   | 116            | 634            | 1192           | 2772                | 38.5           | 36.8           | 29.7           | 50.0           | 49.7             | 48.3           | 44.5           | 51.7           | 11.2           | 11.5           | 14.8           | 1.7            |
| /ʌ/  |       | cuff   | 113            | 641            | 1206           | 2645                | 47.5           | 36.6           | 28.6           | 52.7           | 55.1             | 61.6           | 47.2           | 54.1           | 7.6            | 25.0           | 18.6           | 1.4            |
|      |       | dun    | 113            | 626            | 1235           | 2740                | 43.0           | 28.2           | 32.2           | 57.6           | 48.2             | 50.5           | 48.5           | 50.8           | 5.2            | 22.3           | 16.3           | -6.8           |
|      |       | pup    | 113            | 605            | 1235           | 2645                | 48.9           | 25.9           | 32.1           | 48.9           | 53.0             | 53.6           | 49.4           | 46.1           | 4.1            | 27.7           | 17.3           | -2.8           |
|      |       | sud    | 135            | 605            | 1341           | 2676                | 41.7           | 28.0           | 37.0           | 47.1           | 56.0             | 53.0           | 50.1           | 47.4           | 14.3           | 25.0           | 13.1           | 0.3            |

| Female | Vowel  | Word | Formant (Hz)   |                |                | Stimulus Level (dB) |                |                | Noise Level (dB) |                |                | Relative dB    |                |                |      |      |      |
|--------|--------|------|----------------|----------------|----------------|---------------------|----------------|----------------|------------------|----------------|----------------|----------------|----------------|----------------|------|------|------|
|        |        |      | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub>      | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub>   | F <sub>3</sub> | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub> |      |      |      |
| /f/    | leave  | 235  | 352            | 2809           | 3285           | 27.4                | 48.8           | 51.0           | 50.3             | 44.5           | 48.5           | 52.6           | 49.9           | 17.1           | -0.3 | 1.6  | -0.4 |
|        | peak   | 235  | 358            | 2763           | 3423           | 30.8                | 43.3           | 50.3           | 53.1             | 46.5           | 48.2           | 49.3           | 45.6           | 15.7           | 4.9  | -1.0 | -7.5 |
|        | sheen  | 245  | 364            | 2951           | 3452           | 18.7                | 43.9           | 46.9           | 50.9             | 46.7           | 50.1           | 52.2           | 47.2           | 28.0           | 6.2  | 5.3  | -3.7 |
|        | teethe | 245  | 357            | 2974           | 3463           | 24.1                | 44.9           | 44.0           | 49.5             | 45.5           | 50.5           | 52.2           | 52.6           | 21.4           | 5.6  | 8.2  | 3.1  |
|        | peel   | 248  | 333            | 2708           | 3060           | 27.4                | 48.0           | 47.7           | 53.4             | 50.1           | 50.4           | 48.6           | 49.3           | 22.7           | 2.4  | 0.9  | -4.1 |
| /l/    | wig    | 218  | 433            | 2382           | 3025           | 36.4                | 12.5           | 44.9           | 45.9             | 54.5           | 44.6           | 52.4           | 50.1           | 18.1           | 32.1 | 7.5  | 4.2  |
|        | dig    | 216  | 429            | 2362           | 2786           | 36.9                | 11.2           | 47.1           | 53.5             | 53.5           | 43.3           | 52.6           | 53.9           | 16.6           | 32.1 | 5.5  | 0.4  |
|        | fill   | 216  | 429            | 2362           | 2786           | 36.0                | 29.4           | 52.8           | 44.6             | 58.2           | 45.3           | 51.4           | 49.1           | 22.2           | 15.9 | -1.4 | 4.5  |
|        | pick   | 234  | 456            | 2219           | 2838           | 32.1                | 17.9           | 55.7           | 47.3             | 46.1           | 43.2           | 52.8           | 50.5           | 14.0           | 25.3 | -2.9 | 3.2  |
|        | sing   | 234  | 462            | 2454           | 3009           | 26.0                | 13.5           | 48.2           | 49.3             | 45.6           | 54.0           | 54.9           | 49.9           | 19.6           | 40.5 | 6.7  | 0.6  |
| /e/    | led    | 229  | 670            | 2020           | 2903           | 36.5                | 20.6           | 31.2           | 50.9             | 53.6           | 43.2           | 51.8           | 50.8           | 17.1           | 22.6 | 20.6 | -0.1 |
|        | ten    | 226  | 675            | 2037           | 2951           | 31.7                | 21.7           | 43.8           | 54.0             | 48.1           | 51.7           | 55.0           | 51.4           | 16.4           | 30.0 | 11.2 | -2.6 |
|        | rent   | 218  | 622            | 1891           | 2786           | 37.7                | 36.3           | 37.5           | 47.7             | 50.0           | 46.6           | 49.4           | 51.2           | 12.3           | 10.3 | 11.9 | 3.5  |
|        | nest   | 208  | 634            | 1951           | 2740           | 41.3                | 35.7           | 34.9           | 45.2             | 54.9           | 51.1           | 53.4           | 52.0           | 13.6           | 15.4 | 18.5 | 6.8  |
|        | bat    | 216  | 1061           | 1883           | 2939           | 41.4                | 30.0           | 39.1           | 48.6             | 51.2           | 51.2           | 44.1           | 54.4           | 9.8            | 21.2 | 5.0  | 5.8  |
| /æ/    | lash   | 218  | 1053           | 1907           | 2951           | 41.8                | 22.8           | 28.4           | 51.1             | 51.2           | 51.9           | 50.5           | 53.4           | 9.4            | 29.1 | 22.1 | 2.3  |
|        | mat    | 216  | 1053           | 1907           | 2974           | 38.2                | 32.3           | 40.4           | 48.5             | 52.1           | 49.3           | 51.6           | 50.9           | 13.9           | 17.0 | 11.2 | 2.4  |
|        | pass   | 216  | 1079           | 1938           | 2786           | 38.2                | 26.7           | 35.1           | 45.4             | 50.5           | 46.3           | 54.3           | 52.8           | 12.3           | 19.6 | 19.2 | 7.4  |
|        | hath   | 221  | 1098           | 1997           | 2871           | 34.5                | 23.1           | 32.6           | 52.1             | 47.5           | 48.7           | 51.6           | 48.5           | 13.0           | 25.6 | 19.0 | -3.6 |
|        | dumb   | 208  | 627            | 1252           | 2695           | 37.5                | 29.8           | 22.1           | 55.4             | 50.5           | 49.0           | 50.6           | 49.3           | 13.0           | 19.2 | 28.5 | -6.1 |
| /ʌ/    | cuff   | 216  | 704            | 1292           | 2809           | 38.6                | 32.2           | 30.5           | 52.7             | 52.0           | 52.2           | 51.8           | 53.1           | 13.4           | 20.0 | 21.3 | 0.4  |
|        | dun    | 216  | 643            | 1273           | 2763           | 38.7                | 37.9           | 29.0           | 48.8             | 49.9           | 48.4           | 49.6           | 52.7           | 11.2           | 10.5 | 20.6 | 3.9  |
|        | pup    | 242  | 712            | 1250           | 2838           | 27.9                | 21.9           | 31.7           | 53.1             | 49.6           | 53.6           | 46.6           | 51.7           | 21.7           | 31.7 | 14.9 | -1.4 |
|        | sud    | 221  | 664            | 1325           | 2838           | 39.2                | 24.9           | 34.1           | 50.3             | 46.1           | 48.4           | 51.6           | 53.1           | 6.9            | 23.5 | 17.5 | 2.8  |

## D. CM-in utero

| Male | Vowel | Word   | Formant (Hz)   |                |                | Stimulus Level (dB) |                |                | Noise Level (dB) |                |                | Relative dB    |                |                |      |      |      |      |
|------|-------|--------|----------------|----------------|----------------|---------------------|----------------|----------------|------------------|----------------|----------------|----------------|----------------|----------------|------|------|------|------|
|      |       |        | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub>      | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub>   | F <sub>3</sub> | F <sub>0</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub> |      |      |      |      |
| /i/  | leave | peak   | 124            | 361            | 2409           | 3587                | 34.4           | 31.3           | 43.0             | 46.3           | 45.9           | 44.8           | 47.0           | 46.8           | 11.5 | 13.5 | 4.0  | 0.5  |
|      |       | sheen  | 135            | 357            | 2584           | 3504                | 34.8           | 34.4           | 49.2             | 44.8           | 49.8           | 44.0           | 52.6           | 48.2           | 15.0 | 9.6  | 3.4  | 3.4  |
|      |       | teethe | 126            | 337            | 2524           | 3546                | 33.2           | 34.8           | 42.1             | 48.5           | 55.3           | 46.0           | 47.6           | 46.5           | 22.1 | 11.2 | 5.5  | -2.0 |
|      |       | peel   | 163            | 317            | 2495           | 3716                | 24.7           | 41.6           | 47.6             | 47.5           | 46.0           | 46.1           | 46.7           | 49.9           | 21.3 | 4.5  | -0.9 | 2.4  |
|      |       | wig    | 116            | 374            | 2117           | 2614                | 35.1           | 19.9           | 46.9             | 44.0           | 46.5           | 44.8           | 46.4           | 50.3           | 11.4 | 24.9 | -0.5 | 6.3  |
| /ɪ/  | dig   | fill   | 124            | 406            | 1928           | 2614                | 32.9           | 21.5           | 41.8             | 45.6           | 44.5           | 48.4           | 47.9           | 48.7           | 11.6 | 26.9 | 6.1  | 3.1  |
|      |       | pick   | 127            | 357            | 1997           | 2645                | 35.8           | 27.0           | 47.7             | 50.2           | 46.6           | 42.5           | 45.6           | 50.9           | 10.8 | 15.5 | -2.1 | 0.7  |
|      |       | sing   | 135            | 451            | 2245           | 2838                | 32.4           | 32.3           | 46.5             | 44.0           | 46.6           | 45.5           | 51.1           | 48.7           | 14.2 | 13.2 | 4.6  | 4.7  |
|      |       | led    | 117            | 626            | 1735           | 2805                | 35.0           | 23.8           | 46.7             | 50.2           | 46.1           | 45.3           | 44.2           | 46.1           | 11.1 | 21.5 | -2.5 | -4.1 |
|      |       | ten    | 135            | 591            | 1928           | 2772                | 29.3           | 35.5           | 43.6             | 48.2           | 44.9           | 47.6           | 44.8           | 48.4           | 15.6 | 12.1 | 1.2  | 0.2  |
| /æ/  | rent  | nest   | 124            | 538            | 1755           | 2676                | 28.1           | 23.3           | 45.1             | 50.2           | 49.2           | 46.0           | 45.2           | 46.6           | 21.1 | 22.7 | 0.1  | -3.6 |
|      |       | bat    | 112            | 688            | 1776           | 2437                | 47.7           | 45.2           | 49.5             | 49.0           | 44.6           | 46.8           | 47.9           | 46.9           | -3.1 | 1.6  | -1.6 | -2.1 |
|      |       | lash   | 113            | 696            | 1655           | 2614                | 34.0           | 22.7           | 44.3             | 44.0           | 42.6           | 43.9           | 48.3           | 46.7           | 8.6  | 21.2 | 4.0  | 2.7  |
|      |       | mat    | 115            | 626            | 1735           | 2645                | 35.3           | 30.6           | 47.2             | 52.4           | 48.0           | 47.7           | 50.4           | 50.7           | 12.7 | 17.1 | 3.2  | -1.7 |
|      |       | pass   | 113            | 672            | 1598           | 2614                | 36.6           | 30.3           | 51.0             | 45.8           | 46.9           | 39.5           | 41.6           | 46.8           | 10.3 | 9.2  | -9.4 | 1.0  |
| /ʌ/  | hath  | dumb   | 113            | 649            | 1617           | 2524                | 37.1           | 27.7           | 47.9             | 48.7           | 44.8           | 48.5           | 52.3           | 47.1           | 7.7  | 20.8 | 4.4  | -1.6 |
|      |       | cuff   | 116            | 634            | 1192           | 2772                | 30.3           | 27.4           | 37.2             | 41.4           | 47.6           | 44.2           | 44.8           | 42.8           | 17.3 | 16.8 | 7.6  | 1.4  |
|      |       | dun    | 113            | 641            | 1206           | 2645                | 33.0           | 25.2           | 37.6             | 48.2           | 48.6           | 46.1           | 42.9           | 46.6           | 15.6 | 20.9 | 5.3  | -1.6 |
|      |       | pup    | 113            | 626            | 1235           | 2740                | 32.7           | 27.1           | 39.8             | 46.1           | 48.3           | 43.6           | 47.0           | 44.0           | 15.6 | 16.5 | 7.2  | -2.1 |
|      |       | sud    | 113            | 605            | 1235           | 2645                | 35.2           | 40.3           | 41.8             | 50.0           | 45.3           | 47.6           | 46.2           | 44.3           | 10.1 | 7.3  | 4.4  | -5.7 |
| 135  | 605   | 1341   | 2676           | 30.5           | 43.4           | 51.1                | 48.7           | 42.8           | 42.3             | 46.8           | 50.2           | 12.3           | -1.1           | -4.3           | 1.5  |      |      |      |



## REFERENCES

- Abrams, R. M., Gerhardt, K. J., & Antonelli, P. J. (1998). Fetal hearing. Developmental Psychobiology, 33 (1), 1-3.
- Abrams, R. M., Gerhardt, K. J., Griffiths, S. K., Huang, X., & Antonelli, P. J. (1998). Intrauterine sound in sheep. Journal of Sound and Vibration, 216 (3), 539-542.
- Abrams, R. M., Griffiths, S. K., Huang, X., Sain, J., Langford, G., & Gerhardt, K. J. (1998). Fetal music perception: the role of sound transmission. Music Perception, 15, 307-317.
- Abrams, R. M., Hutchinson, A. A., McTernan, M. J., & Merwin, G. E. (1987). Effects of cochlear ablation on local cerebral glucose utilization in fetal sheep. American Journal of Obstetrics and Gynecology, 157, 1438-1442.
- Arjmand, E., Harris, D. M., & Dallos, P. (1988). Developmental changes in frequency mapping of the gerbil cochlea: comparison of two cochlea locations. Hearing Research, 32, 93-97.
- Armitage, S. E., Baldwin, B. A., & Vince, M. A. (1980). The fetal sound environment of sheep. Science, 208, 1173-1174.
- Aslin, R. N. (1987). Visual and auditory development in infancy. In: Osofsky, J. (Ed.), Handbook of Infant Development (2nd Edition). New York: A Wiley-Interscience Publication, John Wiley & Sons, Inc. 5-97.
- Aslin, R. N., Pisoni, D. B., & Jusczyk, P. W. (1983). Auditory development and speech perception in infancy. In: Haith, M. M., & Campos, J. J. (Eds.), Handbook of Child Psychology, Volume 2: Infancy and Developmental Psychobiology (4th Edition). New York: John Wiley & Sons, Inc. 573-687.
- Bench, J. (1968). Sound transmission to the human foetus through the maternal abdominal wall. The Journal of Genetic Psychology, 113, 85-87.

- Benzanquen, S., Gagnon, R., Hunse, C., & Foreman, J. (1990). The intrauterine sound environment of the human fetus during labor. American Journal of Obstetrics and Gynecology, 163 (2), 484-490.
- Berg, W. K., & Berg, K. M. (1987). Psychophysiological development in infancy: State, startle, and attention. In: Osofsky, J. (Ed.), Handbook of Infant Development (2nd Edition). New York: A Wiley-Interscience Publication, John Wiley & Sons, Inc. 238-317.
- Bilger, R. C., & Wang, M. D. (1976). Consonant confusions in patients with sensorineural hearing loss. Journal of Speech and Hearing Research, 19, 718-748.
- Birnholz, J. C., & Benacerraf, B. R. (1983). The development of the human fetal hearing. Science, 222, 516-518.
- Borden, G. J., & Harris, K. S. (1984). Speech Science Primer: Physiology, Acoustics, and Perception of Speech (2nd Edition). Baltimore, MD: Williams and Wilkins.
- Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Kotby, M. N., Nasser, N. H. A., El Kholy, W. A. H., Nakanish, Y., Oyer, H., Powell, R., Stenphens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G. I., Westerman, S., & Ludvigsen, C. (1994). An international comparison of long-term average speech spectra. The Journal of the Acoustical Society of America, 96, 2108-2120.
- Cooper, R. P., & Aslin, R. N. (1989). The language environment of the young infant: implications for early perceptual development. Canadian Journal of Psychology, 43 (2), 247-265.
- Cox, R. M., & Moore, J. N. (1988). Composite speech spectrum for hearing aid gain prescriptions. Journal of Speech and Hearing Research, 31, 102-107
- Davis, H. (1948). The articulation area and the social adequacy index for hearing. Laryngoscope, 58, 761-778.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: newborns prefer their mothers' voice. Science, 208 (6), 1174-1176.
- DeCasper, A. J., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C., & Maugeais, R. (1994). Fetal reaction to recurrent maternal speech. Infant Behavior and Development, 17, 159-164.
- DeCasper, A. J., & Prescott, P. A. (1984). Human newborns' perception of male voices: preference, discrimination, and reinforcing value. Developmental Psychobiology, 17 (5), 481-491.

- DeCasper, A. J., & Sigafos, A. D. (1983). The intrauterine heartbeat: a potent reinforcer for newborns. Infant Behavior and Development, 6, 19-25.
- DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. Infant Behavior and Development, 9, 133-150.
- Dunn, H. K., & White, S. D. (1940). Statistical measurements on conversational speech. The Journal of the Acoustical Society of America, 11, 278-288.
- Echteler, S. M., Arjmand, E., & Dallos, P. (1989). Developmental alterations in the frequency map of the mammalian cochlea. Nature, 341, 147-149.
- Egan, J. P. (1948). Articulation testing methods. Laryngoscope, 58, 955-991.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. Science, 171, 303-306.
- Fifer, W. P. (1995). Neonatal preference for mother's voice. In: Krasnegor, N. A., Blass, E. M., Hofer, M. A., & Smotherman, W. P. (Eds.), Perinatal Development: A Psychobiological Perspective. Orlando, FL: Academic Press, Inc., Harcourt Brace Jovanovich, Publishers. 111-124.
- Fifer, W. P., & Moon, C. M. (1988). Auditory experience in the fetus. In: Smotherman, W. P., & Robinson, S. R. (Eds.), Behavior of the Fetus. West Caldwell, NJ: Telford Press. 175-188.
- Fifer, W. P., & Moon, C. M. (1989). Psychobiology of newborn auditory preferences. Seminars in Perinatology, 13 (5), 430-433.
- Fifer, W. P., & Moon, C. M. (1994). The role of mother's voice in the organization of brain function in the newborn. Acta Paediatr (Stockholm), Suppl. 397, 86-93.
- Fifer, W. P., & Moon, C. M. (1995). The effects of fetal experience with sound. In: Lecanuet, J.-P., Fifer, W. P., Krasnegor, N. A., & Smotherman, W. P. (Eds.), Fetal Development: A Psychobiological Perspective. New Jersey: Lawrence Erlbaum Associates, Inc. 351-368.
- Fitch, R. H., Miller, S., & Tallal, P. (1997). Neurobiology of speech perception. Annual Review of Neuroscience, 20, 331-351.
- Fletcher, H. (1953). Speech and Hearing in Communication. New York: D. Van Nostrand Company, Inc.
- Fodor, J. A. (1983). The Modularity of the Mind. Cambridge, MA: MIT Press.



- Fowler, C. A. (1996). Listeners do hear sounds, not tongues. The Journal of the Acoustical Society of America, 99, 1730-1741.
- French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sound. The Journal of the Acoustical Society of America, 19, 90-119.
- Gagnon, R., Benzaquen, S., & Hunse, C. (1992). The fetal sound environment during vibroacoustic stimulation in labor: effect on fetal heart rate response. Obstetric & Gynecology, 79, 950-955.
- Gelman, S. R., Wood, S., Spellacy, W. N., & Abrams, R. M. (1982). Fetal movements in response to sound stimulation. American Journal of Obstetrics and Gynecology, 143, 484-485.
- Gerhardt, K. J. (1989). Characteristics of the fetal sheep sound environment. Seminars in Perinatology, 13 (5), 362-370.
- Gerhardt, K. J. (1990). Prenatal and perinatal risks of hearing loss. Seminars in Perinatology, 14 (4), 299-304.
- Gerhardt, K. J., & Abrams, R. M. (1996). Fetal hearing: Characterization of the stimulus and response. Seminars in Perinatology, 20 (1), 11-20.
- Gerhardt, K. J., Abrams, R. M., & Oliver, C. C. (1990). Sound environment of the fetal sheep. American Journal of Obstetrics and Gynecology, 162, 282-287.
- Gerhardt, K. J., Huang, X., Arrington, K. E., Meixner, K., Abrams, R. M., & Antonelli, P. J. (1996). Fetal sheep *in utero* hear through bone conduction. American Journal of Otolaryngology, 17, 374-379.
- Gerhardt, K. J., Otto, R., Abrams, R. M., Colle, J. J., Burchfield, D. J., & Peters, A. J. M. (1992). Cochlear microphonics recorded from fetal and newborn sheep. American Journal of Otolaryngology, 13, 226-233.
- Griffiths, J. D. (1967). Rhyming minimal contrasts: a simplified diagnostic articulation test. The Journal of the Acoustical Society of America, 42, 236-241.
- Griffiths, S. K., Brown, W. S., Gerhardt, K. J., Abrams, R. M., & Morris, R. J. (1994). The perception of speech sounds recorded within the uterus of a pregnant sheep. The Journal of the Acoustical Society of America, 96, 2055-2063.
- Grimwade, J. C., Walker, D. W., & Wood, C. (1970). Sensory stimulation of the human fetus. Australian Journal of Mental Retardation, 2, 63-64.

- Gulick, W. L., Gescheider, G. A., & Frisina, R. D. (1989). Hearing: Physiological Acoustics, Neural Coding, and Psychoacoustics. New York: Oxford University Press, Inc.
- Harris, D. M., & Dallos, P. (1984). Ontogenetic changes in frequency mapping of a mammalian ear. Science, 225, 741-743.
- Hawkins, A. D., & Myrberg, A. A. (1983). Hearing and sound communication under water. In: Lewis, B. (Ed.), Bioacoustics: A Comparative Approach. New York: Academic Press. 347-405.
- Hawkins, J. E., Jr., & Stevens, S. S. (1950). The masking of pure tones and of speech by white noise. The Journal of the Acoustical Society of America, 22, 6-13.
- Hepper, P. G. (1992). Fetal psychology: an embryonic science. In: Nijhuis, J. G. (Ed.), Fetal Behaviour: Developmental and Perinatal Aspects. Oxford: Oxford University Press. 129-156.
- Hepper, P. G., Scott, D., & Shahidullah, S. (1993). Newborn and fetal response to maternal voice. Journal of Reproductive and Infant Psychology, 11, 147-153.
- Hepper, P. G., & Shahidullah, S. B. (1994a). The development of fetal hearing. Fetal and Maternal Medicine Review, 6, 167-179.
- Hepper, P. G., & Shahidullah, S. B. (1994b). Development of fetal hearing. Archives of Disease in Childhood, 71, F81-F87.
- Hillenbrand, J., Getty, L. A., Clark M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. The Journal of the Acoustical Society of America, 97, 3099-3111.
- Hirsh, I. J., Reynolds, E. G., & Joseph, M. (1954). Intelligibility of different speech materials. The Journal of the Acoustical Society of America, 26, 530-538.
- Hollien, H., & Feinstein, S. (1975). Contribution of the external auditory meatus to auditory sensitivity underwater. The Journal of the Acoustical Society of America, 57, 1488-1492.
- Honrubia, V., & Ward, P. H. (1968). Longitudinal distribution of the cochlear microphonics inside the cochlear duct (guinea pig). The Journal of the Acoustical Society of America, 44, 951-958.
- Jackson, H., Hackett, J. T., & Rubel, E. W. (1982). Organization and development of brain stem auditory nuclei in the chicken: ontogeny of postsynaptic responses. Journal of Comparative Neurology, 210, 80-86.

- Johansson, B., Wedenberg, E., & Westén, B. (1964). Measurement of tone response by the human foetus: A preliminary report. Acta Otolaryngologica (Stockholm), 57, 188-192.
- Jusczyk, P. W. (1996). Developmental speech perception. In: Lass, N. J. (Ed.), Principles of Experimental Phonetics. St. Louis, MO: Mosby-Year Book, Inc. 328-361.
- Jusczyk, P. W., Pisoni, D. B., Reed, M. A., Fernald, A., & Myers, M. (1983). Infants' discrimination of the duration of a rapid spectrum change in nonspeech signals. Science, 222, 175-176.
- Kent, R. D. (1997). The Speech Sciences. San Diego, CA: Singular Publishing Group, Inc.
- Kisilevsky, B. S., & Muir, D. W. (1991). Human fetal and subsequent newborn responses to sound and vibration. Infant Behavior and Development, 14, 1-26.
- Kryter, K. D. (1946). Effects of ear protective devices on the intelligibility of speech in noise. The Journal of the Acoustical Society of America, 18, 413-417.
- Kuhl, P. K. (1987). Perception of Speech and Sound in Early Infancy. In: Salapatek, P., & Cohen, L. (Eds.), Handbook of Infant Perception, Volume 2: From Perception to Cognition. New York: Academic Press, Inc. 275-382.
- Kuhl, P. K. (1992). Psychoacoustics and speech perception: internal standards, perceptual anchors, and prototypes. In: Werner, L. A., & Rubel, E. W. (Eds.), Developmental Psychoacoustics. Washington, DC: American Psychological Association. 293-332.
- Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. Science, 190, 69-72.
- Kuhl, P. K., & Miller, J. D. (1978). Speech perception by the chinchilla: identification functions for synthetic VOT stimuli. The Journal of the Acoustical Society of America, 63, 905-917.
- Lecanuet, J.-P. (1996). Prenatal auditory experience. In: Deliege, I., & Sloboda, J. (Eds.), Musical Beginnings: Origins and Development of Musical Competence. Oxford: Oxford University Press. 3-34.
- Lecanuet, J.-P., Granier-Deferre, C., & Busnel, M.-C. (1989). Differential fetal auditory reactivity as a function of stimulus characteristics and state. Seminars in Perinatology, 13 (5), 421-429.

- Lecanuet, J.-P., Granier-Deferre, C., & Busnel, M.-C. (1991). Prenatal familiarization. In: Bonniec, G. P.-L., & Dolitsky, M. (Eds.), Language Bases... Discourse Bases: Some Aspects of Contemporary French-Language Psycholinguistics Research. Amsterdam, Netherlands: John Benjamins Publishing Company. 31-44.
- Lecanuet, J.-P., Granier-Deferre, C., & Busnel, M.-C. (1995). Human fetal auditory perception. In: Lecanuet, J.-P., Fifer, W. P., Krasnegor, N. A., & Smotherman, W. P. (Eds.), Fetal Development: A Psychobiological Perspective. New Jersey: Lawrence Erlbaum Associates, Inc. 239-262.
- Lecanuet, J.-P., Granier-Deferre, C., Jacquet, A.-Y., & Busnel, M.-C. (1992). Decelerative cardiac responsiveness to acoustic stimulation in the near term fetus. Quarterly Journal of Experimental Psychology, 44b, 279-303.
- Lecanuet, J.-P., Granier-Deferre, C., Jacquet, A.-Y., Capponi, I., & Ledru, L. (1993). Prenatal discrimination of a male and a female voice uttering the same sentence. Early Development and Parenting, 2 (4), 217-228.
- Lecanuet, J.-P., & Schaal, B. (1996). Fetal sensory competencies. European Journal of Obstetric & Gynecology and Reproductive Biology, 68, 1-23.
- Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: developmental changes of temporal and spectral parameters. The Journal of the Acoustical Society of America, 105, 1455-1468.
- Liberman, A. M. (1982). On the finding that speech is special. American Psychologist, 37, 148-167.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. Cognition, 21, 1-36.
- Lippe, W. R., & Rubel, E. W. (1983). Development of the place principle: tonotopic organization. Science, 219, 514-516.
- Lippe, W. R., & Rubel, E. W. (1985). Ontogeny of tonotopic organization of brain stem auditory nuclei in the chicken: implications of development of the place principle. Journal of Comparative Neurology, 237, 273-289.
- Marty, R. (1962). Développement postnatal des réponses sensorielles du cortex cérébral chez le lapin. Archives d'Anatomie Microscopique et de Morphologie Expérimentale, 51, 129-264.

- Mehler, J., Bertoncini, J., Barrière, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. Perception, 7, 491-497.
- Mehler, J., & Dupoux, E. (1994). What Infants Know: The New Cognitive Science of Early Development (Translated by Southgate, P.). Cambridge, MA: Blackwell.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. Cognition, 29, 143-178.
- Miller, G. A., & Nicely, P. E. (1955). An Analysis of perceptual confusions among some English consonants. The Journal of the Acoustical Society of America, 27, 338-352.
- Mills, D. M., Norton S. J., & Rubel, E. W. (1994). Development of active and passive mechanics in the mammalian cochlea. Auditory Neuroscience, 1, 77-99.
- Mills, D. M., & Rubel, E. W. (1996). Development of the cochlear amplifier. The Journal of the Acoustical Society of America, 100, 1-15.
- Mills, M., & Melhuish, E. (1974). Recognition of mother's voice in early infancy. Nature, 252 (8), 123-124.
- Moon, C. M., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. Infant Behavior and Development, 16, 495-500.
- Moon, C. M., & Fifer, W. P. (1990). Syllables as signals for 2-day-old infants. Infant Behavior and Development, 13, 377-390.
- Northern, J. L., & Downs, M. P. (1991). Hearing in Children (4th Edition). Baltimore, MD: Williams & Wilkins.
- Ohala, J. J. (1996). Speech perception is hearing sounds, not tongues. The Journal of the Acoustical Society of America, 99, 1718-1725.
- Peck, J. E. (1994). Development of hearing, part II: embryology. Journal of American Academic of Audiology, 5, 359-365.
- Penrod, J. P. (1985). Speech discrimination testing. In: Katz, J. (Ed.), Handbook of Clinical Audiology (3rd Edition). Baltimore, MD: Williams & Wilkins. 235-255.
- Peters, A. J. M., Abrams, R. M., Gerhardt, K. J., & Griffiths, S. K. (1993a). Transmission of airborne sound from 50-20,000 Hz into the abdomen of sheep. Journal of Low Frequency Noise and Vibration, 12, 16-24.

- Peters, A. J. M., Gerhardt, K. J., Abrams, R. M., & Longmate, J. A. (1993b). Three-dimensional intraabdominal sound pressures in sheep produced by airborne stimuli. American Journal of Obstetric and Gynecology, 169, 1304-1315.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. The Journal of the Acoustical Society of America, 24, 175-184.
- Pickles, J. O. (1988). An Introduction to the Physiology of Hearing (2nd Edition). San Diego, CA: Academic Press, Inc.
- Pollack, I. (1948). Effects of high pass and low pass filtering on the intelligibility of speech in noise. The Journal of the Acoustical Society of America, 20, 259-266.
- Pujol, R., & Hilding, D. (1973). Anatomy and physiology of the onset of auditory function. Acta Otolaryngology (Stockholm), 76, 1-11.
- Pujol, R., Lavigne-Rebillard, M., & Lenoir, M. (1998). Development of sensory and neural structures in the mammalian cochlea. In: Rubel, E. W., Popper, A. N., & Fay, R. R. (Eds.), Development of Auditory System. New York: Springer-Verlag New York Inc. 146-192.
- Pujol, R., Lavigne-Rebillard, M., & Uziel, A. (1989). Physiological correlates of development of the human cochlea. Seminars in Perinatology, 14 (4), 275-280.
- Pujol, R., & Uziel, A. (1989). Auditory development: Peripheral aspects. In: Timiras, P.S., & Meisami, E. (Eds.), Handbook of Human Biologic Developmental Biology, Vol.I: Neural, Sensory, Motor, and Integrative Development, Part B: Sensory, Motor, and Integrative Development. Boca Raton, FL: CRC Press. 109-130.
- Querleu, D., Renard, X., Boutteville, C., & Crepin, G. (1989). Hearing by the human fetus? Seminars in Perinatology, 13 (5), 409-420.
- Querleu, D., Renard, X., & Crépin, G. (1981). Perception auditive et réactivité foetale aux stimulations sonores. Journal of Gynecology & Obstetric and Biological Reproduction, 10, 307-314.
- Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crepin, G. (1988a). Fetal hearing. European Journal of Obstetric & Gynecology and Reproductive Biology, 29, 191-212.
- Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Vervoot, P. (1988b). Intra-amniotic transmission of the human voice. Review of French Obstetric & Gynecology, 83 (1), 43-50.

- Richards, D. S., Frentzen, B., Gerhardt, K. J., McCann, M. E., & Abrams, R. M. (1992). Sound levels in the human uterus. Obstetrics & Gynecology, 80, 186-190.
- Romand, R. (1987). Tonotopic evolution during development. Hearing Research, 28, 117-123.
- Romand, R., & Romand, M.-R. (1982). Myelination kinetics of spiral ganglion cells. Journal of Comparative Neurology, 204, 933-942.
- Rosner, B. S., & Doherty, N. E. (1979). The response of neonates to intra-uterine sounds. Developmental Medicine and Child Neurology, 21, 723-729.
- Rubel, E. W. (1978). Ontogeny of structure and function in the vertebrate auditory system. In: Jacobson, M. (Ed.), Handbook of Sensory Physiology, Development of Sensory Systems (Vol. IX). New York: Springer-Verlag Press. 135-237.
- Rubel, E. W. (1984). Ontogeny of auditory system function. Annual Review of Physiology, 46, 213-219.
- Rubel, E. W. (1985a). Auditory system development. In: Gottlieb, G., & Krasnegor, N. A. (Eds.), Measurement of Audition and Vision in the First Year of Postnatal Life: A Methodological Overview. Norwood, NJ: Ablex Publishing Corporation. 53-90.
- Rubel, E. W. (1985b). Strategies and problems for future studies of auditory development. Acta Otolaryngologica (Stockholm), Suppl. 421, 114-128.
- Rubel, E. W., & Ryals, B. M. (1983). Development of the place principle: acoustic trauma. Science, 219, 512-514.
- Rubel, E. W., Smith, D. J., & Miller, L. C. (1976). Organization and development of brain stem auditory nuclei of the chicken: ontogeny of nucleus magnocellularis and nucleus laminaris. Journal of comparative Neurology, 166, 469-490.
- Ruben, R. J. (1992). The ontogeny of human hearing. Acta Otolaryngologica (Stockholm), 112, 192-196.
- Rübsamen, R., & Lippe, W. R. (1998). The development of cochlear function. In: Rubel, E. W., Popper, A. N., & Fay, R. R. (Eds.), Development of Auditory System. New York: Springer-Verlag New York Inc. 193-270.
- Salk, L. (1962). Mother's heartbeat as an imprinting stimulus. Transactions of the New York Academy of Sciences, Series 2 (4), 753-763.

- Schill, H. A. (1985). Threshold for speech. In: Katz, J. (Ed.), Handbook of Clinical Audiology (3rd Edition). Baltimore, MD: Williams & Wilkins. 224-234.
- Schweitzer, L., & Cant, N. B. (1984). Development of the cochlear innervation of the dorsal cochlear nucleus of the hamster. Journal of Comparative Neurology, 225, 228-243.
- Shahidullah, S., & Hepper, P. G. (1993). The developmental origins of fetal responsiveness to an acoustic stimulus. Journal of Reproductive and Infant Psychology, 11, 135-142.
- Shahidullah, S., & Hepper, P. G. (1994). Frequency discrimination by the fetus. Early Human Development, 36, 13-26.
- Smotherman, W. P., & Robinson, S. R. (1995). Tracing developmental trajectories into the prenatal period. In: Lecanuet, J.-P., Fifer, W. P., Krasnegor, N. A., & Smotherman, W. P. (Eds.), Fetal Development: A Psychobiological Perspective. New Jersey: Lawrence Erlbaum Associates, Inc. 15-32.
- Spence, M. J., & DeCasper, A. J. (1987). Prenatal Experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. Infant Behavior and Development, 10, 133-142.
- Starr, A., Amlie, R. N., Martin, W. H., & Sanders, S. (1977). Development of auditory function in newborn infants revealed by auditory brainstem potentials. Pediatrics, 60 (6), 831-839.
- Thornton, A. R., & Raffin, M. J. M. (1978). Speech-discrimination scores modeled as a binomial variable. Journal of Speech and Hearing Research, 21, 507-518.
- Vince, M. A., Armitage, S. E., Baldwin, B. A., Toner, J., & Moore, B. C. J. (1982). The sound environment of the foetal sheep. Behaviour, 81, 296-315.
- Vince, M. A., Billing, A. E., Baldwin, B. A., Toner, J. N., & Weller, C. (1985). Maternal vocalizations and other sounds in the fetal lamb's sound environment. Early Human Development, 11, 179-190.
- Walker, D., Grimwade, J., & Wood, C. (1971). Intrauterine noise: a component of the fetal environment. American Journal of Obstetric and Gynecology, 109, 91-95.
- Walsh, E. J., & McGee, J. (1990). Development of auditory coding in the central nervous system: implications for *in utero* hearing. Seminars in Perinatology, 14 (4), 281-293.



- Wang, M. D. (1976). SINFA: multivariate uncertainty analysis for confusion matrices. Behavior Research Methods and Instrumentation, 8, 471-472.
- Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: a study of perceptual features. The Journal of the Acoustical Society of America, 54, 1248-1266.
- Wang, M. D., Reed, C. M., & Bilger, R. C. (1978). A comparison of the effects of filtering and sensorineural hearing loss on patterns of consonant confusions. Journal of Speech and Hearing Research, 21, 5-36.
- Wilkins, W. K., & Wakefield, J. (1995). Brain evolution and neurolinguistic preconditions. Behavioral and Brain Sciences, 18, 161-226.
- Winer, B. J., Brown, D. R., & Michels, K. M. (1991). Statistical Principles in Experimental Design (3rd Edition). New York: McGraw-Hill, Inc.
- Wollack, C. H. (1963). The auditory activity of the sheep (ovis aries). Journal of Auditory Research, 3, 121-132.
- Yancey, D., & Dallos, P. (1985). Ontogenetic changes in cochlear characteristic frequency at a basal turn location as reflected in the summating potential. Hearing Research, 18, 189-195.
- Yost, W. A. (1994). Fundamentals of Hearing: An Introduction (3rd Edition). San Diego, CA: Academic Press, Inc.

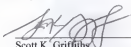
## BIOGRAPHICAL SKETCH

Xinyan Huang was born on July 9, 1964, in Beijing, China, where he was raised and educated. He obtained his M.D. degree from Beijing Medical University in 1988, and started his residency in otolaryngology-head and neck surgery at the Third Teaching Hospital, Beijing Medical University, where he was honored as Doctor of the Year in 1991 and as Teacher of the Year in 1992. Treating patients with sensorineural hearing loss stimulated his interest in studying hearing science. As a result, he came to the United States to pursue his Ph.D. degree after completing his otolaryngology residency in 1993. As a Graduate Research Associate, he joined the Perinatology Research Laboratory at the University of Florida in 1993. Working under the direction of Dr. Kenneth J. Gerhardt and Dr. Robert M. Abrams, he had been involved in several research projects on fetal hearing. In 1998, he was honored to receive the Outstanding Academic Achievement Award by the University of Florida. After completed his graduate study, he will begin his postdoctoral training at San Diego, California. He married Min Feng in 1993. Their son, Alvin Tianyi Huang, was born on June 17, 1999.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

  
Kenneth J. Gerhardt, Chairman  
Professor of Communication Sciences  
and Disorders

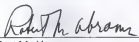
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

  
Scott K. Griffiths  
Associate Professor of Communication  
Sciences and Disorders


I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

  
F. Joseph Kemker  
Professor of Communicative Disorders

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

  
Robert M. Abrams  
Professor of Obstetrics and Gynecology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

  
\_\_\_\_\_  
Kyle E. Rarey  
Professor of Anatomy and Cell Biology

This dissertation was submitted to the Graduate Faculty of the Department of Communication Sciences and Disorders in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1999

\_\_\_\_\_  
Dean, Graduate School

LD  
1780  
1999  
.H874